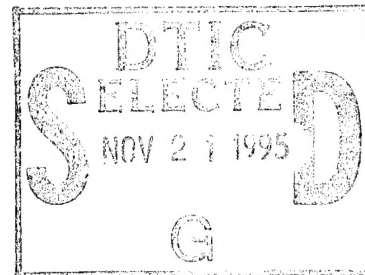


Decision-Making With Long-Term Consequences: Temporal Discounting for Single and Multiple Outcomes in the Future

Mary Kay Stevenson
Purdue Research Foundation



Research and Advanced Concepts Office
Michael Drillings, Acting Chief

October 1995



QUALITY INSPECTED

United States Army
Research Institute for the Behavioral and Social Sciences

19951115 020

U.S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES

**A Field Operating Agency Under the Jurisdiction
of the Deputy Chief of Staff for Personnel**

EDGAR M. JOHNSON
Director

Research accomplished under contract
for the Department of the Army

Purdue Research Foundation

Technical review by

Michael Drillings

Accession For	
NTIS CRA&I	<input checked="checked" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification _____	
By _____	
Distribution /	
Availability Codes	
Dist	Avail and / or Special
A-1	

NOTICES

DISTRIBUTION: This report has been cleared for release to the Defense Technical Information Center (DTIC) to comply with regulatory requirements. It has been given no primary distribution other than to DTIC and will be available only through DTIC or the National Technical Information Service (NTIS).

FINAL DISPOSITION: This report may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The views, opinions, and findings in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other authorized documents.

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503				
1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE 1995, October	3. REPORT TYPE AND DATES COVERED FINAL 8/89 - 7/92		
4. TITLE AND SUBTITLE Decision-Making With Long-Term Consequences: Temporal Discounting for Single and Multiple Outcomes in the Future		5. FUNDING NUMBERS MDA903-98-K-0224 0601102A B74F 2901 C44		
6. AUTHOR(S) Mary Kay Stevenson (Purdue University)				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Purdue Research Foundation Hovde Hall, 3rd Floor West Lafayette, IN 47907		8. PERFORMING ORGANIZATION REPORT NUMBER		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences ATTN: PERI-BR 5001 Eisenhower Ave. Alexandria, VA 22333-5600		10. SPONSORING/MONITORING AGENCY REPORT NUMBER ARI Research Note 96-04		
11. SUPPLEMENTARY NOTES COR: Michael Drillings				
12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.		12b. DISTRIBUTION CODE		
13. ABSTRACT (Maximum 200 words): The consequences of a decision may be characterized as positive or negative, certain or uncertain, and as immediate or deferred. The first two attributes have been studied extensively. The impact of future consequences has received much less attention and is the focus of the current studies. Although temporal discounting is expected, by most normative models, to occur as a function of time, an empirical comparison of the discounting function, applied to singular and multiple outcomes described in this paper, contradicts this expectation in one domain.				
14. SUBJECT TERMS Decision-making Cognitive process Response measures Discounting functions			15. NUMBER OF PAGES 100	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT Unlimited	

DECISION-MAKING WITH LONG-TERM CONSEQUENCES: TEMPORAL DISCOUNTING FOR SINGLE AND MULTIPLE OUTCOMES IN THE FUTURE

CONTENTS

	Page
EXPERIMENT 1	5
Methods	12
Results	16
Positive Outcomes	18
Negative Future Consequences	27
Discussion	32
EXPERIMENT 2	35
Methods	40
Results	43
Discussion	55
REFERENCES	59
FIGURE CAPTIONS	71

LIST OF TABLES

Table 1. The reliability of responses for the individual subjects across four sessions.....	63
2. The subjective values obtained for the individual subjects for the magnitude and delay of the outcomes.....	64
3. The sample of funding program attributes	67
4. The discounting weights estimated for the groups and individuals.....	68
5. Reliability and discounting weights for the individual subjects.....	69

LIST OF FIGURES

Figure 1a. Conceptual model for the evaluation of risky funding programs.....	73
1b. Conceptual model for the comparison of riskfree funding programs.....	73
2. The fit of the multiplicative model for the risky funding program described as positive evaluations.	74
3. The mean preference ratings obtained for the funding programs.....	75

DECISION-MAKING WITH LONG-TERM CONSEQUENCES: TEMPORAL DISCOUNTING FOR SINGLE AND MULTIPLE OUTCOMES IN THE FUTURE

CONTENTS (Continued)

	Page
Figure 4. The form of the response function describing strength of preference for positive outcomes.....	76
5. The scale values for the funding programs estimated for the preference task platted as a function of their attributes: support and time.....	77
6. Psychophysical functions for the amount of funding and work time requirement.....	78
7. The fit of the multiplicative model for the risky funding program described as negative evaluations.	79
8. The form of the response function describing strength of preference for negative outcomes.....	80
9. The scale values for the loan programs estimated from the preference task plotted as a function of their attributes: pay-back period and loan time.....	81
10. Psychophysical functions for the pay-back periods and loan time.....	82
11. An example of a multiple outcome work-study program	83
12a. The response function for Group R describing the fit of the averaging model.....	84
12b. The response function for Group L describing the fit of the averaging model.....	85
13. The observed judgments for work-study programs offering support for two years and single years for Group R (a & b) Adding Strategy (top panels c & d) Averaging Strategy (bottom panels).....	86
14. The observed judgments for work study programs offering support for two years and single years for Group L (a & B) Adding Strategy(top Panel c & d) Averaging Strategy (bottom panel).....	87
15a. Set size variations I the observed ratings for the group classified as average in Group R.....	88
15b. Set size variations I the observed ratings for the group classified as average in Group L.....	89
16a. The response function for Group R describing the fit of the additive model.....	90
16b. The response function for Group L describing the fit of the additive model.....	91
17a. The response function for Group R describing the fit of the additive model.....	92
17b. The response function for Group L describing the fit of the additive model.....	93

DECISION-MAKING WITH LONG-TERM CONSEQUENCES: TEMPORAL DISCOUNTING
FOR SINGLE AND MULTIPLE OUTCOMES IN THE FUTURE

CONTENTS (Continued)

	Page
Figure 18a. Set size variations in the observed ratings for the group classified as additive Group R.....	94
18b. Set size variations in the observed ratings for the group classified as additive Group L.....	95
19a. The response function for Group R describing the fit of the set size model.....	96
19b. The response function for Group L describing the fit of the set size model.....	97

Decision-Making With Long-Term Consequences:

Temporal Discounting for Single and Multiple Outcomes in the Future

In many situations, we are faced with making a decision that involves outcomes that will not materialize until some time in the future. Normative models (Koopmans, 1960; Meyers, 1976) for future outcomes assume that people prefer more immediate gains and prescribe that value of the future outcomes be discounted as a function of their temporal distance. Most temporal discounting models proposed in economics assume: 1. Positive outcomes that are remote in time are less preferred than positive outcomes that are more immediate. 2. Negative outcomes that are delayed are preferred to negative outcomes that occur sooner. 3. Discount rates are constant across time. These models specify that the difference in the temporal factors determines the values of outcomes, not the absolute time when they occur (Benzion, Rapoport, & Yagil, 1989).

Very little empirical work has been done on this type of decision to determine what individuals actually do in a variety of situations when asked to evaluate future outcomes. Mischel and his colleagues (Mischel & Grusec, 1967, Mischel & Metzner, 1962; Miller & Karniol, 1976; Mischel, Grusec & Masters, 1969; Mischel & Ebbeson, 1970; Moore, Mischel & Zeiss, 1976; Yates & Mischel, 1979) have reported a series of studies with children on the factors influencing delay of gratification. Miyamoto & Eraker (1988) have described the characteristics of health care decisions that specify survival duration. Ainslie (1975) has studied impulse control and Stevenson (1986), Loewenstein (1987), Benzion et al, (1989) described the discounting characteristics that subjects exhibited when asked to evaluate investments. However, there have been several published

comments on the lack of empirical work on deferred consequences in decision-making (Bjorkman, 1984; von Winterfeldt & Fischer, 1975; Nisan & Minkowich, 1973, Benzion et al, 1989).

The earliest work on the effects of delaying consequences was done with animals in basic learning theory. These learning models focused on three variables: the magnitude, probability and immediacy of reinforcers. In general, it is a well-established finding that delaying the consequences of a response reduces its impact on behavior (Renner, 1964; Tarpy & Sawabini, 1974). Another well-established finding is that the effect of delay can be attenuated by presenting secondary reinforcement cues during the delay interval (Renner, 1964; Tarpy & Sawabini, 1974). In other words, it is easier to establish an association between a response and a delayed outcome, if cues that have been associated with the delayed reward are presented during the delay interval. The current studies look at the characteristics of temporal discounting for single outcomes and for a series of events. In the spirit of this early work in learning theory, we might expect that when events are described in series, the discounting mechanism might be attenuated for long-term consequences due to the intervening events.

There are several ways of interpreting the reasons why people evaluate future gains as less attractive than immediate gains and future losses as less distressing than immediate losses. Mischel and his colleagues have proposed that subjects find waiting aversive. Therefore, the time interval will detract from the positive outcomes, making them less desirable. However, this interpretation implies that subjects will also find future losses more distressing, since waiting increases the agony. Previous research with loans (Stevenson, 1986)

indicates that, for future payments, this interpretation is not true. This result does not imply that all negative consequences are less distressing as they are put off into the future but does imply that, in some cases, people prefer to delay negative events.

Mischel and his colleagues also proposed that temporal discounting might occur because subjects interpret the delay as probability. In others words, subjects prefer immediate gains because they have learned that some future events never materialize. Therefore, if you must wait, there is a chance that the event will not occur. This interpretation implies that future losses should be less distressing than immediate losses. A criminal sentenced to death would prefer a delayed penalty, in the hopes that, during the delay, an appeal would change the future. This interpretation is consistent with the results obtained by Stevenson (1986). Yates (1975) pointed out that these two interpretations can be distinguished by comparing the relative values of positive and negative outcomes.

However, there is another interpretation that can be used to describe the reason why people discount future events. In some cases, the future event may be certain. For example, you are guaranteed a particular return on an investment, and yet you prefer to invest for a shorter period of time. Temporal discounting may occur because the utility of a future event is unknown. This hypothesis implies that the time interval introduces uncertainty into the decision process. This uncertainty represents the absence of information about the circumstances that will surround that future event. For example, it is not clear, when you invest money, whether or not you will have an emergency that requires the use of the vested money. This type of uncertainty has not been

conceptualized in any utility models about future outcomes. Note that these interpretations may apply to different situations. The criminal sentenced to death has a known utility for future death. So, in this case, a delay is desirable because there is a possibility that the execution will be cancelled. However, there are many situations that indicate the uncertainty of the future may relate to the utility of the outcome rather than to the probability of its occurrence. In summary, temporal discounting is a complex phenomena that represents one of the real-world forms of uncertainty that is incorporated in many decisions. In order to understand the way the future impacts on our decisions, we must broaden our concept of temporal discounting and explore its effects in many situations.

These experiments were designed to test the generality of the ratio discounting model (Stevenson, 1986; Miyamoto & Eraker, 1988; Benzion et al, 1989) by assessing how students think about work study programs to fund their college education when the programs involve either a single outcome or multiple outcomes.

Experiment 1

Students were asked to evaluate financial aid programs that varied in the amount of support that could be earned and the length of time they worked to obtain support. Unlike the investment situation where subjects waited for the outcomes to occur, in this situation, subjects were required to fill the time period with work.

These stimuli are quite similar to the decision made by new army or navy recruits. Military personnel are required to complete a tour of duty before they receive support for college. In two tasks, representing positive future consequences, the students were asked to consider funding programs for college

that required them to work for a period of time in order to obtain support for school. In the two additional tasks, the students evaluated funding programs that require them to repay, after they have finished school, the support they receive for college. This type of funding program is defined as a future negative consequence or loss because the cost factor is delayed. This type of program is very similar to the conditions offered in ROTC programs. There are some differences between the laboratory and applied setting (e.g., ROTC participants are required to complete some training while attending school), however, the temporal structure of the incentives is similar.

With investment stimuli, support has been obtained for the ratio discounting operation (Stevenson, 1986, Benzion, et al, 1989) with a rating task,

$$[1] \quad R_{pmt} = J_r [S_p S_m / S_t].$$

where R_{pmt} is the rated attractiveness of the investment, J_r is a negatively accelerated response function, S_p is the subjective value of the probability of a profit, S_m is the subjective value of the magnitude of the return, and S_t is the subjective value of the time required for the investment to mature.

Support for a ratio discounting function was also obtained with a preference task (Stevenson, 1986) that required the subjects to indicate how much they preferred one investment over another one:

$$[2] \quad P_{AB} = J_p [(S_{mA}/S_{tA}) - (S_{mB}/S_{tB})].$$

where P_{AB} is the strength of preference for one investment over the other (A or B), J_p is an S-shaped response function, S_{mB} is the subjective value of the profit for Investment A, S_{mA} is the subjective value of the profit for Investment B, S_{tA} is the subjective value of the time required for Investment A to mature, and S_{tB} is the subjective value of the time required for Investment

B to mature. Therefore, regardless of the task, the value of future outcomes were reduced proportionately according to their temporal proximity. The larger the value, the greater the reduction in future utility. The discounting operation was consistent across tasks. The first goal of the current study is to determine if a ratio discounting function will be obtained for the funding programs for college.

A conceptual framework for the cognitive processes proposed for these two judgment tasks is shown in Figures 1a and 1b. The subject is presented with the objective description of the funding program. The objective values of the attributes are represented with the ϕ_a . The subject interprets the values of the attributes in terms of their experience and judgment. The subjective values of the attributes (S_a) are then combined in order to arrive at an evaluation of the funding program (∇_{fp}). The discounting operation describes the way the time factor influences the value of the support. This evaluation must be represented on the response scale provided by the experimenter. This process is assumed to be a monotonic mapping from the internal value to the objective response format required by the experimenter. Although this framework is similar to the models described by Anderson (1970; 1982) as functional measurement, different assumptions are made in this conceptualization. For example, the response function is not assumed to be linear; the form of the response function is defined with basis splines so that the best descriptive form can be obtained. Furthermore, the additive and multiplicative models which predict the same rank order for positive stimuli are distinguished by using probability as a third multiplicative factor in the rating task or a two operation model (Birnbaum & Veit, 1974) in the preference task. Although several theorists (eg., Birnbaum,

1974; Shanteau, 1975; Mellers, 1982; and others) have applied this type of model to a wide range of judgment problems, different assumptions and analytical techniques have been used to define the characteristics of these components.

Insert Figures 1a and 1b about here.

The same measurement model has been used in Figure 1b to describe the preference task. In this case, subjects will be asked to compare two funding opportunities and indicate the degree to which they prefer one or the other. When two funding programs are presented, the attributes of each stimulus are viewed subjectively. The subject may combine the attributes of each stimulus in order to arrive at an overall value for each alternative (V_A and V_B) and then compare these values. A subtractive operation is usually used to represent the comparison process in a preference task (Winsberg and Ramsay, 1981). The preference value ($V_A - V_B$) is then mapped onto the response scale provided by the experimenter by the function J_p .

The second issue for this study concerns the scale values of the attributes of the stimuli being judged. Consider the model diagrammed in Figure 1a. The subjects were given the following information:

Work in the community for 2 years and
Receive 75% of your college expenses.
50% of the students flunk out of this college.

In order to determine if the ratio discounting operation was used, a probability factor had to be added to the problem. The students were told that the participants would lose the support if they flunked out of school (similar to the GI benefits). Previous research indicates that probability and value

combine multiplicatively; so, given that assumption, three discounting models that predict different rank orders for the funding programs can be discriminated:

$$[3a] \quad R_{pmt} = J_r [S_p S_m / S_t]$$

$$[3b] \quad R_{pmt} = J_r [S_p (S_m - S_t)]$$

$$[3c] \quad R_{pmt} = J_r [S_p S_m - S_t]$$

The first model indicates that time affects the value of the support proportional to its original value, while the last two models specify that the effect of time is independent of the original value.

Recall (see Figure 1a) that the subjects are assumed to translate the physical values describing the funding program: probability (ϕ_p), magnitude of support (ϕ_m), and time spent working (ϕ_t) into subjective values (S_p , S_m , S_t , respectively). Initially, the scale values for each attribute were assumed to be independent of the other attributes present; however, the previous research with investments (Stevenson, 1986) indicated that the subjective scale values varied across situations that include and exclude explicit probabilities. It is not clear whether or not this failure to obtain scale convergence across tasks using risky and riskfree stimuli will also occur when students are evaluating funding programs for school.

This effect of probability on the values of a return has been discussed by Schoemaker (1982). Simply stated, value functions (i.e., subjective values derived in a riskfree context) are only monotonically related to utilities (i.e., subjective values that are obtained in the context of risk). These implications are straightforward. Gains and losses are valued differently in riskfree and risky contexts. In other words, incentive programs to enlist will be evaluated differently, depending on the recruit's perception of the risk involved in

obtaining the benefits by joining the service. In the previous study, a return on an investment that included a risk factor was evaluated as greater than the same return from a riskfree investment. The inverse was true for losses. Certain loss was evaluated as more negative than a loss that was uncertain. A similar pattern was obtained for the scale values for time in the two tasks. Subjective time was consistently evaluated as longer, when waiting for a riskfree positive investment return, and shorter, when waiting to pay off a loan. In this experiment, the same subjects responded to positive and negative future outcomes and risky and riskfree future consequences. Therefore, the relative values of discounted outcomes can be compared across gains and losses under risky and riskfree contexts. According to economic theory (Schoemaker, 1981), the value of the support should vary with the presence and absence of risk. Given the previous study by Stevenson (1986), the effect of time should also be affected by the presence and absence of risk.

The third component of the discounting model describes the way the subjects use the response scale. After the subjects combine the subjective values of the attributes or discount the value of the outcome according to the risk and temporal factors, they must express their judgment on the response scale provided by the experimenter. The response function describes this mapping strategy and is represented by J_r in the rating model. A single pole rating scale has been consistently associated with a negatively accelerated response function. This means that the subjects tend to scale the most desirable investments more similarly than less desirable investments. When losses are introduced, subjects tend to reverse the strategy so that the most undesirable investments are rated more similarly than less undesirable losses. The same response strategy should

be represented in the current study since the form of the function is hypothesized to depend on the task and not on the attributes of the stimulus set.

An S-shaped response function was obtained in the previous studies with the bipolar preference ratings (eg., Winsberg & Ramsay, 1979; Rose & Birnbaum, 1975). This function indicated that subjects tended to rate stimuli as more similar when preference is very strong, near the ends of the scale (i.e., away from indifference). This pattern is similar to the results obtained with the rating scale since, in both cases, the extreme stimuli are rated as more similar to each other than is predicted by the ratio discounting model. Both forms of these response functions should be replicated in the current study.

When two stimuli are presented, the subject may also compare each attribute and then generate a preference based on the comparison of the individual attributes. This model will also be tested :

$$[4] \quad P_{mt} = J_p [(S_{mA} - S_{mB}) - (S_{tA} - S_{tB})].$$

where the parameters are defined in the same way as Equation 2. The previous findings with investment stimuli indicated that subjects used a wholist strategy in comparing the investments in the preference task as opposed to an attribute comparison process (eg. Tversky, 1969).

In summary, the objectives of this experiment are primarily concerned with extending the discounting model to a different stimulus set. First, it will be important to establish the generality of the ratio discounting function in situations that require an intervening activity during the time interval. Secondly, the form of the response function can be tested with the same tasks but with different stimuli to verify the interpretation of the response characteristics of preference ratings and evaluations. Finally, the relationships

established in the previous studies between the subjective scale values in the riskless and risky stimulus sets for gains and losses can be replicated in this new context.

Methods

Each subject was instructed to act as a financial aid counselor who had to advise students who could not afford to attend college without financial assistance. Each task involves a different type of funding arrangement. All of the subjects rated each of the funding programs described below. One set of programs required the student to work for a period of time in order to earn a percentage of the school expenses. In another set of programs, the possibility of flunking out of school was included in the decision scenario. In this case, participants could lose the support that they had earned, so there was a degree of explicit uncertainty about the future outcome. Both of these programs describe a positive incentive system in which the student provides a service and is given support in return. In the third set of programs, the student was permitted to attend school for various periods of time and then was required to forfeit 25% of his or her income over a period of time in order to repay the loan. In the fourth set of stimuli, the students attend school and then a certain percentage of the participants are required to repay the loan, introducing an element of explicit uncertainty into the program. The last two tasks described negative outcomes that are deferred to the future.

Risky Work-Study Programs. Each subject was asked to rate the desirability of a set of work-study programs that varied in the work time requirement, the amount of support that could be earned, and in the probability the student would flunk out of the school and forfeit the support. The continuous rating scale was

defined on the left side, with the worst deal, and on the right, with the best deal.

Sixty-four stimuli were constructed from a factorial combination of four work time requirements (6 months, 1 year, 2 years, 4 years), four degrees of support (25%, 50%, 75%, 100%) and the probability of flunking out of school and losing the support (.75, .50, .25, 0).

Work-Study Preference Task. Subjects rated their degree of preference for one of two work-study programs presented on each trial. The work-study programs varied in the length of time the student would be required to work before attending the college of his or her choice. These programs also varied in the amount of support (e.g., a percentage of the expenses for tuition, room, board and books). Subjects compared each set of work-study programs and indicated their preference for one of the work-study conditions by moving the cursor to the appropriate location on a continuous response scale. The center of the response scale represented indifference. Moving to the left or right indicated the degree of preference for the stimulus on the left or right, respectively. The range of work requirements and possible levels of support were provided at the bottom of the screen so that the subjects could anticipate the worst and best conditions that would occur in the stimulus set.

Sixteen work-study conditions were constructed for Stimulus Set A by a factorial combination of four work time requirements (6 months, 1 year, 2 years, 4 years) and four support levels (25%, 50%, 75%, 100%). The four stimuli of Set B were constructed from a factorial combination of two work time requirements (9 months, 3 years) and two levels of support (33%, 67%). Each of the 16 work-study

plans of Stimulus Set A was combined with each of the four work-study plans for Stimulus Set B, yielding 64 preference trials.

Risky Educational Loans. Each subject was asked to rate the desirability of a set of loan programs that varied in the length of time that support could be obtained, the length of time the participants would be required to repay 25% of their income, and the probability that they would be released from the pay-back requirement. The continuous rating scale was labeled from the worst deal, on the left, to the best deal, on the right. Unlike the rating scale used in the Risky Work-Study Program, the range of time that support could be obtained across all the stimuli, the highest and lowest percentage of students that would be required to repay the loan, and the length of time that they would be required to forfeit 25% of their income was given on the bottom of the screen.

Sixty-four stimuli were constructed from a factorial combination of four levels of support or time to attend college (2 years, 4 years, 6 years, 8 years), four pay back periods or the length of time payments (25% of salary) would be required (1 year, 3 years, 4 years, 6 years), and four percentages describing the probability of actually having to repay the loan (25%, 50%, 75%, 100%).

Educational Loan Preference Task. Subjects rated their degree of preference for one of two educational loan programs presented on each trial. The conditions of the loan varied in the length of time that the participants could attend school and in the length of time they would be required to pay 25% of their income to repay the expenses. Subjects were told that they could attend the college of their choice. The continuous response scale that was described for the Work-Study Preference Task was used to obtain the preference responses. The worst deal was defined by the shortest length of time that could be spent in

school and the longest period of time that would be required to repay the loan and was shown at the bottom of the screen. The best deal was defined by the longest period of time that could be spent in school and the shortest pay-back period and was also shown on the bottom of the screen.

Sixteen loans were constructed for Stimulus Set A by a factorial combination of four support periods (2 years, 4 years, 6 years, 8 years) and four pay-back periods (1 year, 3 years, 4 years, 6 years). Participants were required to forfeit 25% of their salary during the pay-back periods. The four stimuli for Stimulus Set B were constructed from a factorial combination of two support periods (3 years, 6 years) and two pay-back periods (2 years, 5 years). Each of the 16 loans described as Stimulus Set A was combined with the four loans from Stimulus Set B yielding 64 preference trials.

Procedure. Subjects were instructed individually and given several practice trials to check their understanding of each task. Two tasks were completed in a one-hour session. Half of the subjects judged the positive incentive stimuli (Work-Study Programs) during the first session and rated the negative outcome stimuli (Education Loan Programs) one week later, during the second session. The other half of the subjects completed the task in reversed order. Half of the subjects in each condition completed the preference task first, within the session, and the other half completed the rating task first. Both conditions were counterbalanced across subjects.

All of the stimuli were presented individually on video screens and subjects responded on a computer. The cursor was positioned in the center of both the preference scale and the rating scale. Subjects moved the cursor to a position that represented their response and pressed "/". A mark appeared on the scale

shown on the screen. If the subject was satisfied with this location, a second input response recorded the location of the mark and initiated the next trial. Otherwise, the subject was free to move to a new location and repeat the response procedure.

Five subjects were asked to return for three replications of all four tasks and were paid \$4.50 for each additional session. A different pseudo random sequence of trials was used for each subject and replication. Reliability and the characteristics of individual data will be assessed using this subset of the sample.

After the subjects had completed both (or all) of the sessions, they were asked to recall the "best" and "worst" deals from each task and to describe the decision strategy they had used for each type of funding program.

Subjects. Forty-eight undergraduates from Introductory Psychology volunteered to participate for course credit. Seven subjects were used as pilot subjects to test the instructions and were eliminated from the analysis. There were seven subjects who generated patterns of responses that were distinctly different than those generated by the rest of the subjects. These subjects were eliminated from the group analysis which was based on 34 subjects. Five subjects returned to complete 4 replications of each task. These subjects were analyzed individually.

Results

A correlation matrix for each task was computed to compare the pattern of responses that subjects made in order to determine if their responses should be averaged. Each correlation matrix was factored to determine if the response patterns were similar. Most of the subjects loaded on a single factor for each

task. A few subjects appeared to be unique in their approach to one of the four tasks. They were eliminated from the group analysis that was based on the means of the responses made. Therefore, thirty-four subjects were similar in their response patterns (obtained high loadings on a single factor) and were averaged to complete the following analyses. The group analysis for all four tasks was based on the means of the same group of subjects so the resulting parameters could be compared.

The individual subjects were selected prior to any preliminary analyses. The purpose of including the individual analyses is to determine if the results, based on means, for the group corresponds to the results that would be obtained for individual subjects. The responses obtained across four sessions, using the same stimuli but different sequences, were first correlated to determine if the subjects were consistent across sessions. The mean correlation across sessions for each subject and each task is shown in Table 1. The mean response for each subject was obtained across sessions for each task. The percentage of variance in these means predicted by the models described for the group is also listed.

Insert Table 1 about here.

Each task will be described as it relates to the discounting model, first in relation to the group analysis, and then the individual analyses. Then the results obtained with the tasks will be compared.

Positive Outcomes

Risky Funding Programs. The results of an ANOVA indicated that the main effects for each attribute were significant (time: $F(3,99)=175.70$, $p<.05$; support: $F(3,99)=302.01$, $p<.05$; probability of flunking: $F(3,99) = 224.05$, $p<.05$). The three-way interaction among time, support and probability of flunking was also significant ($F(27,891) = 1.75$, $p<.05$) and provides some support for a multiplicative discounting function for time and probability with amount. Therefore, if a linear response function is assumed for these data, the multiplicative model is supported.

The following analyses were used to compare the ordinal consistency of the multiplicative (Equation 3a), distributive (Equation 3b), and dual distributive (Equation 3c) discounting models with the observed responses. The response function was assumed to be monotonic and nonlinearly related to the implicit ratings (V_{pvt}) in order to maximize the fit. First, an average rating was computed for each stimulus across subjects who had similar response patterns. The group analysis was computed on these means. The integrated spline approach (Winsberg & Ramsay, 1981) was adapted to this data to parameterize the response function without making any assumptions about the form of the function. Spline functions are piecewise polynomial functions of a given degree (deBoor, 1978; Winsberg & Ramsay, 1981). The segments of the function are defined by the location of the knots. The knot sequence is used to increase the flexibility of the polynomial function. The internal knot locations will be graphed as vertical lines in all of the appropriate figures. The number of parameters that are required to describe the form of the nonlinear function are determined by the number of knots. The order of the spline function (K) defines the degree of the

polynomial (K-1) for the function. The combined scale values generated by each model (Ψ_{pmt}) were mapped onto a second-order spline function having four segments. Six parameters (including the intercept) were required to describe the response function. Each model was tested using the same number of parameters and order for the splines function, although the placement of the knots varied according to the form of the individual response functions. The response function parameters, four scale values for time, four scale values for the amount of funding, and two probability scale values (the highest and lowest probability of flunking were fixed at .75 and 0, respectively) were estimated using Marquardt's compromise procedure (Draper & Smith, 1981).

The multiplicative model was most accurate in describing the values of the risky funding programs. A plot of the predicted and observed values for this model is shown in Figure 2. The predicted values (Ψ_{fp}) from the model are plotted on the abscissa and the observed responses are represented on the ordinate. The points are coded for the time variable and represent the observed ratings for each investment. The line represents the ratings predicted by a ratio discounting model. The percentage of variance in the ratings accounted for by the distributive model (96.80%) and the dual distributive model (96.56%) were high, but predictions generated from the multiplicative model (99.00%) were most consistent with the observed data. Relatively large deviations were characteristic of the distributive and dual distributive models that were not evident for the multiplicative model. The deviations obtained in fitting these models cannot be reduced by changing the form of the response function, but represent discrepancies in the rank orders of the predicted values and the observed responses. Finally the shape of the response function is negatively

accelerated in agreement with the type of response function obtained in previous studies using investment stimuli (Stevenson, 1986).

Insert Figure 2 about here.

The data obtained from the five subjects who replicated the task was analyzed to determine if the characteristics of the solution obtained for the group or averaged data would also be representative of the individual subjects. The parameters were estimated for each subject for the multiplicative model since it was the best representation of the grouped data. For one subject, there were substantial deviations in the fit so the other models were also fitted to this data. The results of the analysis of the individual data indicated that the multiplicative model was the best representation of the responses in all cases; therefore, subjects used a ratio discounting function for time. The response functions for the individual subjects are similar in form to the response function describing the grouped data. In every case, a negatively accelerated response function optimized the fits. Therefore, the form of the response function obtained for the grouped data appears to represent the response functions of the individual subjects.

In summary, the results obtained for the rating task, using funding programs as stimuli, replicate the results obtained with investments (Stevenson, 1986). The discounting function for time is multiplicative and the response function is negatively accelerated. Furthermore, the results obtained with group data clearly represented the characteristics of the responses made on the individual level.

Therefore, the group means appear to represent the strategy of the individual subjects.

Preference Task. The mean preference ratings, averaged across subjects, obtained for pairs of funding programs, are shown in Figure 3 as points. The predicted values represented by the lines will be described later. The funding programs that were presented from Set A were ordered according to the marginal means of the preference ratings and are represented on the abscissa. Each line (or symbol) represents a different Funding Program from Set B. If subtraction represents the operation used by the subjects to compare the two funding programs, the lines in Figure 3 would be parallel. The barrel-shaped relationship between the lines indicates that either the difference operation does not correspond to the pattern of preference ratings obtained or the response function relating the implicit preference ratings (ρ_{AB}) to the observed responses (P_{AB}) is nonlinear. In order to determine how well the difference model would represent the data when a nonlinear response function is used, an integrated second-order spline function was used to parameterize the response function (J_p) as described for the rating task.

Insert Figure 3 about here.

The predicted values from the difference preference model ($V_A - V_B$) were mapped onto a second-order spline function having four segments to define the form of the function. Therefore the model for the discounting preference model included six parameters (including an intercept) for the response function, 14

scale values for the Funding Programs in Set A (the best and worst funding program were fixed to define the measurement scale), and four scale values for the funding programs in Set B. The parameters were estimated, so as to minimize the sum of the squared deviations of the predicted and observed values, using Marquardt's compromise procedure (Draper & Smith, 1981).

The results of this analysis indicated that the J_p function, maximizing the fit of the difference model, is approximately S-shaped. The form of this function is shown in Figure 4. The observed preference ratings are coded according to the Funding Stimulus represented from Set B and are shown as points. The predicted responses are represented by the line. The discounting preference model (Equation 2) accounted for 98.29% of the variance in the preference ratings. The relationship between the predicted and observed points is also shown in Figure 3. There do not appear to be any large or systematic discrepancies in the fit of this model.

Insert Figure 4 about here.

A model assuming an attribute-by-attribute comparison process was also fitted to the data for contrast. According to the process represented by this model, preference ratings are determined by an additive combination of attribute values. The attribute values are obtained by comparing each characteristic of the set of funding programs. In this case, the amount of work required would be compared across the two funding programs and then the amount of funding available would be compared. This information would then be combined to produce a preference rating between the alternatives. Since time delay represents a

negative attribute, it is assumed to subtract value from the combined evaluation. J_p^* is a monotonic nonlinear response function that will be used to maximize the fit of the attribute comparison model.

In order to complete this analysis, the matrix of preference means was reorganized to include eight combinations of funding rates and eight combinations of work time. In this analysis, six parameters (including the intercept) were used to represent the response function; six scale values for the funding rates (the highest and lowest difference were set equal to constants to define the scale unit), and eight scale values for the time combinations were estimated. The attribute comparison model accounted for 97.76% of the variance in the preference ratings. When the residuals obtained with the attribute comparison model are compared to the residuals obtained with the discounting preference model, it is clear that there is a wider range of discrepancies with the attribute comparison model. There was no way to improve the fit of the attribute comparison model by increasing the flexibility of the response function. The difference in the quality of fit is due to a difference in the rank orders predicted by the two models. Therefore, it appears that the subjects were discounting the funding level for each alternative according to the length of time they would be required to work and then comparing the subjective values of the funding programs.

Discounting Function for support and work time. The estimated scale values obtained from the discounting preference model (Equation 2) for the Funding Programs, Set A and B, are plotted in Figure 5. The scale values for the Funding Level are represented on the abscissa with a different line for each work time requirement. The scale values for each magnitude and time delay are the marginal means of the Funding programs after a linear transformation that will be

described later. The diverging, bilinear fan pattern indicates that a ratio discounting operation describes the discounting function for time on the value of the support offered. For example, this result indicates that the subjects were evaluating the amount of support to be gained according to the length of time that work was required. The same diverging pattern was obtained for both sets of funding programs. This result further supports the discounting preference model and provides evidence against the attribute comparison model. The attribute comparison model defined in Equation 4 requires a different rank order in the stimuli than does a difference of ratio model (Equation 2). If the lines had been parallel in Figure 5, the attribute comparison model would be supported.

Insert Figure 5 about here.

Scale value estimates. Scale values were obtained for the levels of support and work time requirements in the analysis of the risky funding programs and from the preference task that paired riskless funding programs. The marginal means of the scale values for the funding programs, obtained from the discounting preference model, were linearly transformed to match the minimum and maximum estimated scale values obtained from the multiplicative model for the risky funding programs. The psychophysical functions, for both tasks, for the level of support and time to work are shown in Figure 6. For college funding programs, there was a tendency for the scale values for funding from the riskless programs to exceed the scale values for funding obtained from the risky programs. This result contradicts the pattern of scale values obtained across risky and riskless

tasks with investment stimuli. Furthermore, unlike the scale values obtained with the investments, the scale values for work time tend to converge. Therefore, the lack of scale convergence was not replicated with work college funding programs in the degree and form obtained with investment stimuli (Stevenson, 1986). It appears that the change in this relationship between the scale values is due to the stimuli. In this case, the probabilistic attribute was associated with flunking out of school and losing the support. Since flunking depends on performance and is not a random event, the subjects may not have perceived the probability as a risk. If explicit uncertainty or risk determines the adjustment in scale values, its absence in this case could account for the apparent scale convergence for these attributes.

Insert Figure 6 about here.

In summary, the results obtained with the grouped data replicates the discounting operation obtained in the previous studies with investments. Furthermore, the response function is S-shaped, in agreement with the previous research. However, a change in the nature of the probabilistic contingency has been associated with a tendency toward scale convergence in this study.

Since the discounting preference model best represented the group data, that model was fitted to each of the individual subject's data. This model consistently accounted for most of the variance in preference ratings. For individual subjects, the percentage of variance ranged from 96.22% to 99.13%. The preference judgment function was S-shaped for each individual subject although the slope varied across individuals.

For each subject, the scale values obtained for the Funding Programs from Set A and Set B were organized as a function of the scale values for the amount of funding and the time required to work. The patterns obtained for four of the five subjects were almost parallel. The other subject(#27) had diverging lines very similar to the group means. This result indicates that there were individual differences in the pattern of subjective scale values obtained and that there is a need to evaluate individual subjects carefully. Although several general methods (i.e., correlation, factor analysis) were used to evaluate the similarity in strategies across subjects, these methods do not appear to be sensitive enough to discriminate among these models. A method of classifying subjects must be developed. It should be noted that since the grouped analysis indicated a ratio model was most accurate in accounting for the pattern in responses, some portion of the subjects were using that strategy, but it is not clear how many subjects, only that the ratio discounting function dominated in this sample.

The subjective scale values obtained for support and time from the rating and preference tasks were organized for comparison. A linear transformation was done on the scale values obtained from the preference task to adjust the scale proximities on the number line. The relationship described for the group data corresponds very closely to the individual scale values. There is a slight tendency for the scale values, for support generated from the preference task, to exceed the scale values obtained with the rating task. This pattern is opposite to the ordering obtained with investments (Stevenson, 1986). The scale values for time appear to be quite close for each subject, corresponding to the group data.

Negative Future Consequences

Risky Funding Programs. The mean ratings obtained for the set of risky funding programs that required a percentage of the students to repay the loans was obtained across the 37 subjects. If a linear response function is assumed, the results of an ANOVA indicated that there were highly significant main effects for time ($F(3,99) = 183.21, p < .01$), payment ($F(3,99) = 193.71, p < .01$), and probability ($F(3,99) = 121.54, p < .01$). However, none of the interactions were significant. This result contradicts the discounting operations obtained with positive outcomes.

The response function was then allowed to be nonlinear and monotonic and the multiplicative (Equation 1), distributive, and dual distributive discounting models were tested. In each case, 16 parameters were used to fit the data. Five parameters were used to describe the response function. One value represented the intercept; four scale values represented times in school; four scale values represented payment schedules, and two scale values represented probabilities (the lowest and highest values were fixed). As in the case of the positive stimuli, the percentage of variance in the ratings, corresponding to the distributive discounting rule (96.64%) and dual distributive rule (96.81%), is high but not nearly as good as the fit of the multiplicative discounting model (99.15%). More important, however, is the graphical comparison of the fits. The multiplicative model not only provides the closest fit, but there are no systematic deviations in this fit. A plot of the predicted and observed responses for the multiplicative model is shown in Figure 7¹. The graph shows the predicted and observed values of the funding programs, plotted as a function of the implicit value (V_{plt}), derived from the ratio discounting model. The

points are coded for the time spent in school or the delay to the payments. The observed values are those obtained from the mean response of the group. The negatively accelerated response function replicates the form of the function obtained with the positive stimuli as well as the rating judgment function obtained with the investment stimuli. These results also replicate the shift in reference point on the response scale (see Footnote 1).

Insert Figure 7 about here.

The data obtained for the individual comparisons was analyzed, using the multiplicative model, to determine how well the data could be represented. For four subjects, the best fit that could be obtained was with the multiplicative model. The percentage of variance accounted for ranged from 95.39% to 96.81%. A slightly better fit was obtained for one subject with the dual distributive model (92.30% of explained variance), but the irregularity in the response function indicates that this subject may not have been consistent in responding to the program attributes. The response function was very irregular, indicating that the strategy used varied across the scale. For four of the subjects, the response function is negatively accelerated, indicating that the group solution represents these individual response function characteristics.

In summary, the results obtained with delayed negative consequences correspond well with the results obtained with positive consequences and with the investment stimuli. The characteristics of the group analyses appear to be representative of most of the individual subjects.

Preference Task with Negative Outcomes. The mean preference rating across subjects was obtained for pairs of funding programs with different payment schedules.

The predicted values from the discounting preference model (Equation 2) were mapped onto a second order spline function having four segments to define the form of the response function. The parameters for the discounting preference model included six values (including an intercept) for the response function, 14 scale values for the funding programs in Set A (the worst and best funding programs was set to fix the scale unit), and four scale values for the funding programs from Set B. The parameters were estimated so as to minimize the sum of the squared deviations of the predicted and observed values using the same procedure described for the other tasks.

The results of this analysis indicated that the J_p function, maximizing the fit of the difference model, is approximately S-shaped. The response function is shown in Figure 8. The observed values are coded according to stimulus Set B and shown as points. The discounting preference model accounts for 99.61% of the variance in the preference ratings. There are no systematic deviations in the fit of the model.

Insert Figure 8 about here.

The matrix of preference means was rearranged to represent the attribute comparison model as described in the analysis of the preference ratings with positive outcomes. Then the attribute comparison model represented in Equation 4 was fitted for comparative purposes. This model accounts for less of the total

variance, and systematic deviations in the fit are present as you move away from the inflexion point in either direction.

Discounting Function for time in school and amount of pay. The estimated scale values obtained from the discounting preference model (Equation 2) for the Funding Programs in Set A and Set B are plotted in Figure 9. The scale values for the amount of time spent in school (delay of payment) are shown on the abscissa. Each line represents a different payment requirement. The longer the time, the greater the amount to pay. The scale values for these attributes are the marginal means for the scale values of the funding programs after a linear transformation that will be described later. The diverging, bilinear fan pattern indicates that a ratio discounting model describes the discounting operation for time on the value of the payment due.

Insert Figure 9 about here.

Scale value estimates. Scale values were obtained for the length of time spent in school and the length of time required to repay the loan from the riskless preference task and the risky rating task. In this case, the risk factor is more similar to the risk described for the investments when scales did not converge (Stevenson, 1986). The marginal means for the funding programs obtained from the discounting preference model were linearly transformed to match the minimum and maximum estimated scale values obtained with the multiplicative model for the risky funding programs. The psychophysical functions for the stimuli from both tasks are shown in Figure 10. There was a consistent tendency for the riskless payments to be perceived as more costly than the uncertain

payments. This result corresponds with the lack of scale convergence obtained with loan stimuli. The psychophysical functions for time are less clear cut. The slight tendency for the time delay, in the risky funding program, to have larger scale values than the scale values for time delays that were generated with certain outcomes corresponds with the previous results with loans; however, the effect was not as consistent or as pronounced with the student loans. In summary, the results obtained with the preference task, comparing funding programs with delayed payment requirements, replicate the results obtained with positive outcomes and the loan stimuli (Stevenson, 1986), although the scale values for time tend to converge more closely for the funding loans.

Insert Figure 10 about here.

The data obtained from the individual subjects was analyzed using the preference discounting function. The same number of parameters used to fit the group means was used to fit each subject's data. An S-shaped function was consistently obtained across subjects. The discounting preference model accounted for most of the variance ranging from 98.14% to 99.14%. The analysis of the preference ratings obtained from individual subjects clearly indicated that the characteristics of the discounting process and preference task, represented in the means of the group, corresponded to the patterns obtained from individual subjects.

For each subject, the scale values obtained for the Funding Programs from Set A and Set B were organized as a function of the scale values for time spent in school and time spent repaying the loan. The ratio discounting function is

consistently evident in these patterns. The spread in values for the shortest payment duration is longer than the spread in values across time spent in school for the longest pay-back period.

The subjective values for each level of pay-back time and time spent in school is given for each subject in Table 2. The relationship between the scale values for the pay-back period varies across subjects. Some subjects have scale values from the risky task that exceed the scale value from the riskless program and sometimes the relationship is reversed. The inconsistency obtained for the scale values for time in school is not surprising, since the scale values for the group are close together. These results indicate that the effect of risk on the scale values should be investigated further.

Insert Table 2 about here.

Discussion

The purpose of this study was to establish converging evidence for a ratio discounting model using both positive and negative outcomes. The current study extended the previous work to a new type of stimulus. Support for a ratio discounting model was obtained for the grouped data, using funding programs for school that involved working for a period of time and then receiving support and funding programs that permitted the participants to attend school and then repay the support by paying 25% of their income for various periods. The individual data indicated quite clearly that the results described for groups, who had similar response patterns for the negative outcomes, were quite accurate in describing the individual discounting strategy. However, the discounting

strategy with positive outcomes was less consistent. Four of the five subjects (out of 34) appeared to be using a subtractive discounting function and the other subject was quite consistent with the grouped data. Since these subjects only represent a small sample, it is quite likely that most of the subjects could be represented by a ratio discounting function. However, these results do point to the necessity of looking at individual response patterns and the need to develop a better technique of classifying subjects by strategy.

Several aspects of the judgment process were replicated as well. An S-shaped response function was always obtained when the subjects were asked to make preference ratings. A negatively accelerated response function was obtained whenever the subjects were required to rate stimuli on a single directional scale. These response functions were representative of the group means and individual responses. The form of these responses replicates the previous findings with investment stimuli (Stevenson, 1986) and other types of preference stimuli (Winsberg & Ramsay, 1981; Rose & Birnbaum, 1975). These response patterns can be interpreted as the same psychological strategy. In both cases, the subjects have selected a reference point. On the preference scale, it is the center of the response scale. On the rating scale, it is the left or right anchor, depending on whether they are judging gains (left) or losses (right). Subjects appear to be quite adept at scaling their responses near to the chosen reference point. Many stimuli are located in the vicinity of the reference point. However, as the subject moves out on the scale for the extremely good or extremely bad stimuli, he or she rates them as more similar than the attributes would indicate. Therefore, the response function is flat near the end points of the scales.

Finally, the issue of scale convergence was also assessed. Previous results with investment stimuli indicated quite clearly that the scale values for time and value did not converge when estimated for the same subjects from a rating task that included risk and a preference task that described riskless stimuli. For the positive outcome funding programs, when students were told that they would work for a period of time and then be given a percentage of their college expenses, the scale values for time and the scale values for the support were more similar, and almost converged across tasks. In this case, the risky element referred to flunking out of school. The students were told that if they flunked out of school (different probabilities were given), the support would be forfeited even though they had completed the work requirement. This condition is similar to the contingency represented in the military programs. In this case, flunking out might not be perceived as the same type of risk that is represented by a random event. In other words, the student has some control over this event, and, regardless of the reputation of the school, if a particular individual works hard, he or she should be able to determine this likelihood. This result implies that the presence of uncontrollable risk determines whether or not the scale values for the other attributes will converge. This interpretation is consistent with the views described by Schoemaker (1982), who describes the difference between value functions and utility functions that differ in the absence or presence of risk.

For the funding programs that described attending school and then forfeiting 25% of their income for different pay-back periods, scales did not converge as well. In this case, a lottery was incorporated in the risky task. According to the conditions of the lottery, students would be randomly selected and excused

from having to pay back the money used for school. Clearly, this type of risk represented a random event and was out of the control of the students. Although the amount to be paid back looked more similar to the results obtained with loan interest rates, the loan time scales seem to converge. The inconsistency of the individual scale convergence tests indicates that the effect of risk on the scale values for value and time should be investigated more carefully and on an individual level.

In summary, most of the results obtained with funding programs correspond well with the model generated with investment stimuli. Certain features of the data indicate that it is important to assess individual differences in strategy and to consider all the attributes used to describe the temporal outcomes. The model is effectively representing a good portion of the data and is quite consistent for single future outcomes. The next step in the program is to assess the impact of time on the value of a sequence of outcomes that are expected in the future.

Experiment 2

This experiment was designed to determine how subjects would discount a series of events in the future. Subjects were asked to evaluate the desirability of work-study programs that required full time work in the summer and 20 hours of work during the school year in exchange for a percentage of their tuition, room, and board expenses. The work-study programs varied in the number of years that were available in the program and in the amount of support that was available each year. For example, one program might only provide support for the last two years of school (Years 3 and 4). Students

signed up for the program when they entered college and then they began working and receiving support during their third and fourth years. The amount of support varied from 10% to 90%. Therefore, although the work requirement was constant, they could obtain different amounts of support depending on what was available.

The normative discounting model proposed by Meyers (1976) for multiple events is given in Equation 5:

$$[5] \quad y_i = \sum_{i=1}^n \frac{x_i}{(1+r)^i + 1}$$

where y_i is the net present value of an event stream (or series of events x_1, x_2, \dots, x_n) calculated with a constant discount rate, r , over n events or periods, $r > 0$. This model specifies that each event in the series is discounted as a function of time.

According to the work in basic learning theory (Renner, 1964; Tarpay & Sawabini, 1974), using a series of events attenuates the affect of delayed reinforcement. Rats respond as if the "internal clock" is reset by the intervening events whether they are secondary cues or reinforcers. Therefore, one might expect that by providing subjects with support that occurs each year, the events may be discounted at a different and less drastic rate than were they presented as solitary outcomes.

The work-study programs described for this study may be compared to the funding programs described for the first experiment. In the first experiment the students were required to work for a period of time and then they were permitted to attend college with support. Only at the end of the work period

were they provided with the support. In the multiple outcome sets, the students are provided with support over a number of years and are asked to evaluate the program as a whole. A conceptual diagram of this multiple outcome work-study program is shown in Figure 1b. As with single outcomes, the subjects must interpret the value of the support or gains ($S_g = f(\phi_g)$) in relation to their own experience or value system. In order to arrive at an overall value (V_{ws}) of the support offered in the program, each event must be weighted (w_t) for the time when it occurs and then the subjective values of support are combined. For example, if the students use a temporal discounting function, then the weights should decrease as a function of the temporal distance to the year of support (i.e., $w_4 < w_3 < w_2 < w_1$). If the student has the attitude: the more support available, the better the work-study program, then he or she may add the subjective values of the support to arrive at an overall value for the program. If the student believes that the amount of work should reflect the overall gain, the subject may average the support levels that are available for the work time required. In the latter case, the overall value of the work-study program might decrease when an additional year of support is added, if the level of support does not justify the amount of work required. If the student is simply looking at the total support, adding any amount should raise the value of the work-study program.

Finally, the subject must indicate his or her evaluation of the program (V_{wsp}) on the rating scale provided by the experimenter. Since a single pole rating scale was used, we would expect the form of the response function to be negatively accelerated.

In summary, in the single event discounting situation, one event is delayed or evaluated as a future consequence. In the multiple outcome discounting situation, a series of events are anticipated in the future. The discounting operation is represented by the weights. Converging evidence for the discounting function may be obtained with this new situation. A constant discount rate, regardless of the number of intervening events, would be predicted by Meyer's model. Or, the effect of the intervening events may change the subject's tendency to discount more distant events. This effect would correspond to the behavioral effects reported in the basic learning literature on delayed reward.

In order to estimate the weights associated with time, the combination rule for multiple events, the subjective values for the levels of support, and the response function parameters must be estimated.

Two combination functions will be tested, an adding model and an averaging model. The adding model represents the point of view that the more support that can be obtained the better the student likes the program:

$$[6] \quad R_{wsp} = J_r \left[\sum_{i=1}^n w_i S_i \right]$$

where, R_{wsp} is the observed rating, J_r is the monotonic response function for single pole rating scales, w_i is the discounting weight associated with a given year, S_i is the support available for that year, and n is the number of outcomes ($n=4$). It is assumed that, if support is not available for a given year, the w_i parameter is 0. Regardless of the actual support level, this model predicts that adding any support makes the program more attractive. Note that weights can be estimated for an adding model only when the scale

values for the attributes to be combined are identical. If the scale values vary across the attributes, the weights can not be determined in the context of an adding model.

The relative weight averaging model represents the point of view that the value of the program depends on the average value of the support:

$$[7] \quad R_{WSP} = J_r \left[\frac{\sum_{i=0}^n w_i S_i}{\sum_{i=0}^n w_i} \right]$$

where R_{WSP} is the observed rating, J_r is the monotonic response function for single pole rating scales, w_i is the discounting weight associated with a given year, S_i is the support available for that year, and n is the number of outcomes (4). It is assumed that, if support is not available for a given year, the w parameter is 0. S_0 represents the subject's initial expectation for a work study program and w_0 represents the weight the subject associates with that initial impression. If you work the same hours, regardless of the level of support, adding an additional year of support that is relatively low compared to other support levels would reduce the overall value of the program. The students using an averaging model would be expressing a concern that the amount of support obtained across the funded years be worth the work requirement. In other words, they are evaluating the programs by determining the average support per year. In this simple form, this model assumes that the discounting weights are determined by the time of the support and are independent of the number of intervening years of support.

Methods

Each subject was instructed to act as a financial aid counselor who would advise students who could not afford to attend college without financial assistance. For half of the subjects, the two sessions involved rating work study programs and a tuition lottery. For the other half of the subjects, the two sessions involved rating work study programs and student loan programs. Therefore, all of the subjects rated the work study programs described in this report. Due to the complexity of the results, only the results obtained with the work study programs will be reported here. However, all of the tasks will be described.

Work Study Programs. All of the subjects rated the attractiveness of a series of work study programs for college. An example stimulus is shown in Figure 11. Each work study program described the support that could be obtained for one, two, three or four years of college. The students were required to work, full time during the summer and part time during the year(s) that support was available, for minimum wage. In addition, they would receive the amount of support described on the screen. The support includes the expenses needed for tuition, room, board, and books. They did not have to work during the years that support was not available.

Insert Figure 11 about here.

The stimuli selected for evaluation were determined by the combinations summarized in Table 3. The matrix indicates, with x, which support values were available for each year. A total of 187 trials were constructed from

options given in the table. Sixty-four trials described four years of support. All possible combinations of support levels for Year 1, Year 2 (the 50% level was omitted), Year 3, and Year 4 were rated. Forty-eight stimuli described support that was available for 3 years. These stimuli were generated from all possible combinations of support for Years 1, 2, and 3. Sixty-two stimuli described support that was available for two years. All possible two way combinations of Years (1 & 2, 1 & 3, 1 & 4, 2 & 3, 2 & 4, and 3 & 4) were presented for evaluation. Finally, 13 work study programs described only one year of support. All of the marked (x) events in Table 1 were presented for evaluation.

Insert Table 4 about here.

In order to discriminate the adding and averaging models, it was necessary to provide a different number of events in the series. The estimates for the relative weights for each year are determined from the changes in ratings that you get when support for that particular year is not available. In some studies, this design is associated with "missing" information, and the final interpretation of the results depends on whether subjects attempt to estimate the missing information when it is not explicitly provided (Birnbaum, 1980). In this case, there was no reason to rationalize that the year(s) that support was not available were "missing" information. The subjects were told that the availability of support was determined by the investment returns that were financing the work study programs. Subjects who

inferred 0% as the "missing" support would produce a systematic pattern in their ratings, and, therefore, they could be identified.

Variable Interest Rate Loans. For this task, subjects rated the desirability of various student loans that could be obtained to finance their college expenses. Each loan varied in the length of time after graduation that could be used to repay the loan and the interest rate schedule for those years. The results obtained with this task will be described in another report.

Tuition Lottery. For this task, the subjects were asked to rate the attractiveness of various lotteries that they could play to win a percentage of their college expenses. Each lottery represented various combinations of prizes. Each prize was a level of support for school. This task was included in order to compare the effects of probability with time as attributes of outcomes. However, the results are complex and will be described in another report.

Procedure. Each subject worked on two tasks. All of the subjects rated the work-study programs. Half of the subjects rated the loans, and the remaining half rated the lotteries. Each rating task took approximately one hour and was completed with one week intervening. Task order was counterbalanced across subjects.

At the start of the session, subjects were instructed individually about characteristics of the stimuli, rating response, and computer. The cursor was positioned in the center of the rating scale on each trial. Subjects moved the cursor to the location on the scale that represented their evaluation and pressed the slash "/" key. A mark appeared on the rating scale. If the

subject wanted to adjust her response, she could do so as many times as she wanted by moving the cursor and pressing the slash key. If the subject was satisfied with her response, then a second response in the same location ended the trial. The response was stored and the next stimulus was presented.

Subjects completed ten practice trials on the computer before the task was started. During this time, they could ask the experimenter any questions that they might have. After the second session was completed, the subject was debriefed and interviewed about the tasks.

Subjects. Forty-five undergraduate students from an introductory psychology class volunteered for course credit. Twenty-three subjects completed the work study ratings and the lottery ratings (Group R) and twenty-two subjects completed the work study ratings and the loan ratings (Group L). The results will be described, for the ratings of the work study programs, for these groups separately, because of their experience with the other task.

Seven subjects were asked to return for a replication of each task they performed. They were paid \$4.50 for these two sessions. A different trial sequence was used for each subject and replication. Reliability and the characteristics of individual subjects can be assessed using this sample of subjects.

Results

A correlation matrix was computed to compare the pattern of responses obtained across subjects, in order to determine if the characteristics of the responses were similar. The correlation matrix was factored to assess the

similarity of the subjects. Most of the subjects in both groups loaded onto a single factor.

Six points were selected from the 187 ratings, in order to assess whether the individual subjects appeared to be using an adding or averaging strategy. The ratings obtained for work study programs offering support for Year 2 and Year 4 were plotted for each subject, for all 4 combinations of 10% and 90% support (i.e., 10-10, 10-90, 90-10, 90-90). These lines are expected to be parallel for both a simple averaging and an adding strategy. The ratings obtained for the work study program that offered support for Year 2 alone at 10% and 90% was plotted with the two year programs. If this line crossed the two year programs, the subject was classified as averaging. (See Figure 13 - bottom panels for an example of this pattern). If the line for the Year 2 support alone ran parallel to the two year programs, the subject was classified as adding (See Figure 13 - top panels for an example of this type of pattern).

The following analyses were used to compare the ordinal consistency of the observed ratings with a relative weight averaging (Equation 7) and an adding model (Equation 6). The response functions are assumed to be monotonic and may be nonlinearly related to the implicit ratings (Ψ_{WSP}) when the fit is maximized. First an average rating was computed for each work study program across subjects who had been classified as using the same strategy. The integrated spline approach (Winsberg and Ramsay, 1981) was adapted to this analysis to parameterize the response function without making any assumptions about the form of the response function. See Experiment 1 for a short description of spline functions. The predicted values for the Work Study

Programs generated by each model (Ψ_{WSP}) were mapped onto a second-order spline function having four segments. Six parameters were required to describe the response function including the intercept.

Averaging Model. Ten subjects from Group R and 13 subjects from Group L were classified as using an averaging strategy. The mean rating for each Work Study Program from these subjects was computed for each group. The response function parameters (six), four scale values for the levels of support (two scales values were fixed to define the unit), and four weight parameters (one was fixed to define the scale) were estimated using Marquardt's Compromise procedure (Draper & Smith, 1981).

The predicted and observed responses for Group R are shown in Figure 12a and for Group L in Figure 12b. Each graph shows the observed and predicted ratings, plotted as a function of the implicit values (Ψ_{WSP}) of the work study programs generated by the averaging model. The observed points are coded according to the number of years support was available. For Group R, 95.33% of the variance was accounted for with the averaging model. For Group L, 97.20% of the variance was accounted for with the averaging model. There are two important characteristics of these plots to note. First, in both cases, the response function appears to be linear. The negatively accelerated response function, replicated so frequently with single pole rating scales, was not obtained with these work study programs. One possible explanation for the difference in response function patterns is the overall values of these study programs. Since the subjects are combining outcomes, high levels of support are often offset by less extreme support values, and the worst outcomes are often combined with better support values. Perhaps we have

obtained a linear response function because the work-study programs were viewed as exceptional. Second, there appears to be a systematic discrepancy for the single outcome ratings, especially for Group R. These deviations indicate that the model is doing well for outcomes greater than two, but for the single events, the subjects rated them higher than the model predicted.

Several graphs were made to assess the consistency of averaging in the data set. Figure 13 shows the mean ratings of Group R, and Figure 14 shows comparable graphs for Group L. Figures 13c and 14c show the mean ratings for the work study programs offering support for Years 1 and 2 as solid lines. The dashed line shows the mean ratings for the work study programs featuring Year 1 alone at various levels of support. The crossover interaction cannot be modeled using an adding model, regardless of the response function, because the order of the ratings differ from the order predicted by the adding model. Figures 13d and 14d show the work study programs offering support for Years 2 and 3 coded as solid lines. The dashed line represents the mean ratings observed for support that was only available for Year 2. Crossover interactions are present in both cases. Therefore, when the support available for Year 3 is low and combined with the 90% level of support available for Year 2, the attractiveness rating drops as predicted by an averaging process. This crossover interaction was found for all the two year-single year combinations.

Insert Figures 13 & 14 about here.

Figures 15a (Group R) and 15b (Group L) show another pattern that is indicative of an averaging process. The mean ratings for Year 1 support alone is labeled as 1 in this plot. The line labeled 2 shows the mean ratings for work study programs that varied in the level of support for Year 1 but offered 10% support for Year 2. The line labeled 3 shows the mean ratings for work study programs that had various support levels for Year 1 and 10% support available for Years 2 and 3. Finally, the last line, labeled 4, shows the mean ratings for work study programs that had various support levels for Year 1 and 10% support available for Years 2, 3 and 4. Note that, in each case, the mean rating drops as the average support falls. The difference in the lines depends on the scale values for these support values. It is important to note that if subjects had been using a 0 support value for years that support was not available in the work study program, these lines would have to be parallel. It is commonly found that, when subjects "average" information, the more events that are added at the low end of the scale, the greater the rating is. This is termed the set size effect and is represented in the relative weight averaging model with the initial impression parameters, s_0 and w_0 .

 Insert Figure 15 about here.

The weight parameters estimated with the relative weight averaging model for Group R and Group L are listed in Table 4. For Group R, the weight associated with the Year 1 support was larger than the other weights which did not differ from each other. For Group L, the weights were very similar in

value, indicating that subjects did not discount the value of the support as a function of time. This result was quite unexpected. The lack of discounting could be attributed to several interpretations. First, the lack of discounting could be due to the sequential nature of the stimuli. When events are specified in a series, subjects chain the outcomes and combine the set without discounting. Second, the lack of discounting could be due to the type of stimulus situation. Since the problem involved a fixed period of time for the educational process, four years in college, perhaps the subjects did not view the support available for each year as delayed outcomes but as parts to the funding problem for school. Or, discounting could have occurred for some individuals and be hidden by the group average. To assess this possibility, the results obtained for the individual subjects are included in the table. Even for individual subjects, the weights associated with time are nearly identical. In any event, it is quite clear by these results that subjects did not weight the support that was available according to the time.

Insert Table 4 about here.

Since there were systematic deviations in the fit of the relative weight averaging model corresponding to the number of events, another model was tested to determine if those deviations could be reduced. Since the weight parameters estimated were so similar, it is unlikely that they were needed. A Set Size Model was tested that eliminated the weights for the individual years and added weights for the total number of years of support available:

$$[8] \quad R_{WSP} = J_r \left[\frac{w_0 S_0 + v_k \sum_{i=1}^k S_i}{w_0 + v_k} \right]$$

where R_{WSP} is the observed rating for the work study programs, J_r is the response function for rating scales, v_k is a weight associated with the number of supported years available (k varies from 1 to 4), S_i is the scale values for the support levels, w_0 and S_0 are the initial impression parameters.

The response function for the Set Size Model for Group R is shown in Figure 16a and for Group L in Figure 16b. A better fit was obtained for Group R relative to the averaging discounting model. 96.82% of the variance was accounted for with the Set Size model, and no systematic deviations were evident. Furthermore, the weights for the number of outcomes differed across the conditions. The values decreased as a function of the number of events: $v_1=.68$, $v_2=.28$, $v_3=.20$ and $v_4=.16$. Due to the discrepancy in the number of stimuli in each event category, it is difficult to interpret the weights. In general, it appears that single events were rated somewhat higher than multiple events when the support levels were similar. 97.42% of the variance in ratings for Group L was accounted for with the Set Size Model. The weights for the number of outcomes show the same decreasing pattern as the number of years support was available increases: $v_1=1.12$, $v_2=.50$, $v_3=.37$ and $v_4=.30$. The work study programs describing support for only one year tend to be rated higher than those describing support for more than one year when there were similar average support levels. This pattern could be explained by a difference between the set sizes of the amount of work required. When only 1

year is supported, only 1 year of work is required. When support is available for two years, the second year provides the same support level or less in most cases. If support is available for three years, the second and third years provide the same support or less in most cases. Although the absolute amount of support is greater, there may be a tendency for students to expect compensation to increase with experience on the job. Since this does not occur, larger set sizes are viewed as slightly less attractive. This effect is relatively small, but systematic.

Insert Figure 16 about here.

Adding Model Thirteen subjects from Group R and nine subjects from Group L were classified as using an adding strategy. The mean rating for each Work Study Program from these subjects was computed for each group. The response function parameters (six), three scale values for the adding model (two were fixed to define the scale unit) for the levels of support that were available, and three weight parameters for the adding model (one weight was fixed to define the weight scale) were estimated using Marquardt's compromise procedure (Draper & Smith, 1981).

The predicted and observed responses for Group R are shown in Figure 17a and for Group L in Figure 17b. Each graph shows the observed and predicted ratings, plotted as a function of the implicit values (Ψ_{usp}) of the work study programs, generated by the adding model weighted for time. The observed values are shown as points and coded for the number of years support was available. For Group R, 92.06% of the variance in ratings was accounted for

with the discounting additive model. For Group L, 92.53% of the variance was accounted for with the discounting additive model. Consistent with the response function obtained with the averaging strategy, the response functions were nearly linear for both groups. The absence of extreme stimuli could account for this type of function and appears more feasible since the results are consistent across judgment strategies. There also appear to be systematic patterns in the responses for this group, quite different than those obtained for the averaging strategy. In this case, you can see the relationship between the number of events and the mean rating. As the number of years with support increases, the mean rating increases for both groups. Since the deviations for one and two events tend to fall below the predictions of the discounting model and the deviations for three and four years of support tend to fall above the line, it is clear that the Set Size weighting might accommodate variations in the responding that is not modeled well with the discounting model.

Insert Figure 17 about here.

Several graphs were made to assess the consistency of the additive response patterns through the data set. It might be useful to compare these top and bottom panels of Figures 13 and 14 for a contrast of these two models. Figures 13a and 13b show the mean ratings for Group R and Figures 14a and 14b show comparable graphs for Group L. Figures 13a and 14a show the mean ratings for the work study programs offering support for Years 1 and 2 as solid lines. The dashed line shows the mean ratings for the work study programs featuring

Year 1 alone at various levels of support. In both cases, the one year of support is less valued than two years of support, regardless of the level of the support. In other words, even 10% support for the second year is viewed as an improvement. There is no cross over interaction and so the order of the ratings implied by this pattern is consistent with an adding model. Figures 13b and 14b show the work study programs offering support for Years 2 and 3 coded as solid lines. The dashed line represents the mean ratings, observed for support that was only available for Year 2. The rated value for the single year of support is less than the values obtained for the support programs that have two years of support available for both groups. This pattern is consistent with the adding model. The same pattern was obtained for these two groups looking at the other combinations of years.

Figure 18a (Group R) and Figure 18b (Group L) show another pattern that is indicative of the additive process. The mean ratings for Year 1 support alone is labeled 1 on this plot. The line labeled 2 shows the mean ratings for the work study programs that varied in the level of support for Year 1 but offered 10% support for Year 2. The line labeled 3 shows the mean ratings for work study programs that had various support levels for Year 1 and 10% available for Years 2 and 3. Finally, the last line, labeled 4, shows the mean ratings for work study programs that had various support levels for Year 1 and 10% support available for Years 2, 3 and 4. Note that, in each case, the mean rating increases as the additional support is added to the work study program. Consider the rank order of the work study programs representing 90% support for Year 1, and those involving the addition of Year 2-10%, Year 2 and 3-10% and Years 2, 3 and 4-10%. The order obtained with an adding strategy is

the exact opposite of the order obtained for the averaging strategy (See Figure 15a and 15b). If the subjects were discounting for the Year that support was available, we would expect the distance between the lines in the Set Size graphs to decrease as later years are added. This does not seem to be the case.

 Insert Figure 18 about here.

The weight parameters estimated with the discounting additive model for Group R and Group L are listed in Table 4. For both groups, the weights are nearly identical for the four years. This result replicates the results obtained with the averaging model. It is quite clear that the subjects did not discount the value of the support according to the year that it was available, regardless of the combination strategy.

Since there were systematic deviations in the fit of the discounting additive model, corresponding to the number of years support was available, the Set Size Model modification that was applied to the relative weight averaging model was adapted to the additive model:

$$[9] \quad R_{WSP} = J_r \left[v_k \sum_{i=0}^k S_i \right]$$

where R_{WSP} are the observed ratings for the work study programs, J_r is the response function for rating scales, v_k is a set of weights associated with the number of supported years that are available, k varies from 1 to 4, S_i represent the scale values for support levels. Note that the discounting weights have been eliminated from this model.

The response function for the Set Size Model for Group R is shown in Figure 19a and for Group L in Figure 19b. There is a substantial improvement in fit for both groups. 98.15% of the variance in the ratings can be accounted for with the Set Size Model for Group R and 97.85% of the variance in ratings can be described with the Set Size Model for Group L. Furthermore, there do not appear to be any systematic deviations in the fits for either group. However, for Group R, the Set Size weights were quite similar, $v_1=.11$, $v_2=.12$, $v_3=.12$ and $v_4=.10$ and, for Group L, the Set Size weights were also quite similar: $v_1=.10$, $v_2=.12$, $v_3=.13$ and $v_4=.13$. These parameter patterns indicate that a simple additive model could describe the response patterns and the improvement in fit might be attributed to better starting values.

 Insert Figure 19 about here.

Individual Analyses. Seven subjects replicated their responses on each task, so that we could look at the characteristics of the individual subjects and at the reliability of the responses. Table 5 includes the correlations for each subject between sessions for the ratings of the work study task. The values range from .89 to .96. Responses were averaged across sessions. These subjects had been classified for strategy and one subject exhibited patterns similar to the averaging process, while the remaining six subjects exhibited patterns associated with the additive process.

The discounting relative weight averaging model (Equation 7) and the set size averaging model (Equation 8) were tested with the responses obtained from Subject 11. The discounting additive model (Equation 6) and the set size

adding model (Equation 9) were tested for the remaining subjects (10, 26, 57, 61, 62, 65). Table 5 summarizes these results. In every case there was more variability in the weights estimated for set size than for the temporal discounting weights which were nearly equal. It appears that, even on the individual level, there was no discounting as a function of time. The hypothesis that discounting did not occur for the groups because it was averaged over individuals does not appear to be feasible. Further research will be needed to determine if the college situation that is viewed as a block of time or if chaining events together eliminates the discounting effect.

Insert Table 5 about here.

Discussion

The purpose of these studies was to extend the temporal discounting model from single to multiple outcomes. The results were quite surprising. First, discounting was obtained for the funding programs for school that represented single outcomes in the first study. The results were quite similar to the discounting functions obtained with investments. However, there was no discounting for the years of support in the second study. Subjects combined the scale values for support across the years that it was made available, without weighting the support according to the temporal proximity of the support. This effect was replicated across groups and for the individual subjects. Several explanations are possible. First, the discounting effect may not have occurred because the subjects view college as a single 4-year event. In that case, the point in time that support is

obtained would be quite arbitrary, since support is needed to complete the four-year program. An alternative possibility is that the chaining of events may reduce or eliminate the effect of time delays. Chaining has been used in animal studies to facilitate acquisition under long delays of reward (Tarpy & Sawabini, 1974). In order to determine which of these explanations is more feasible, a study needs to be done that uses another type of open-ended situation. The result of this study clearly contradicts the normative model proposed by Meyers (1976).

The response functions obtained were very nearly linear with a single pole rating scale. This result is inconsistent with previous research (Stevenson, 1986, 1987) that consistently obtained a negatively accelerated response function for single outcomes. If the psychological interpretation of the nonlinear functions is applied to the present results, we would say that the implicit values of the stimuli were much less extreme in this study than in the single outcome studies. This is possible, since the current stimuli involved combinations of outcomes. The most favorable situation is often combined with less favorable outcomes and the least favorable situation is often combined with more favorable events. This combination over series of events tends to produce a more mediocre stimulus set. This explanation implies that if the extreme values are eliminated from the investment in funding programs described for single outcomes, the response function would be more linear in character.

Finally, we found that two strategies were represented in this data set. Both the averaging and adding strategies occurred in both groups. An averaging process in this context would mean that the subjects were concerned

about the average support that would be obtained for the amount of work that was required. From this point of view, the low levels of support made the work study program less attractive. An additive process in this context would mean that the student was interested in any support that was possible. From this point of view, even the low levels of support increased the value of the work study program. These results indicate quite clearly that it is important to consider individual differences in modeling the discounting process. Further research is needed to develop more accurate methods of preclassification so that the strategies represented by group analyses represent consistency for the subjects who are averaged.

In conclusion, these studies indicate quite clearly that temporal discounting is a complex phenomena that depends on the circumstances under investigation. Research on the impact of combining events across time appears to be one manipulation that effects the characteristics of the tendency to discount future outcomes.

Footnotes

This research was supported by grant MDA-903-85-K-0366 from the Army Research Institute. The author wishes to express her appreciation to Cindy White and Eric Gallenkamp for assisting with the subjects and preliminary data analyses. I would also like to recognize the contribution made by In Jae Myung for his careful work in assisting in the spline analysis and Marc Brown who prepared some of the graphs.

Requests for reprints should be sent to M. K. Stevenson, Psychological Sciences, Purdue University, West Lafayette, IN, 47907.

1. Note that the response scale has been reversed for these analyses. Previous research (see Stevenson, 1986) found that when subjects are asked to evaluate investments that vary in the amount of money that can be lost, subjects will use the positive anchor and evaluate how "bad" a given contingency is, relative to the positive anchor. By reversing the response scale, the anchor for the rating is on the left, so it corresponds with the anchor subjects use with positive outcomes.

References

- Ainslie, G. (1975). Specious reward: A behavioral theory of impulsiveness and impulse control. Psychological Bulletin, 82, 463-496.
- Anderson, N. H. (1970). Functional measurement and psychophysical judgment. Psychological Review, 77, 153-170.
- Anderson, N. H. (1982). Methods of Information Integration Theory. New York: Academic Press.
- Benzion, U., Rapoport, A., & Yagil, J. (1989). Discount rates inferred from decisions: An experimental study. Management Science, 35(3), 270-284.
- Birnbaum, M. H. (1974). The nonadditivity of personality impressions. Journal of Experimental Psychology, 102, 543-561.
- Birnbaum, M. H. (1980). A recipe for weights and scale values in multiattribute judgment. Unpublished manuscript.
- Bjorkman, M. (1984). Decision making, risk taking and psychological time: Review of empirical findings and psychological theory. Scandinavian Journal of Psychology, 25, 31-49.
- deBoor, C. (1978). A Practical Guide to Splines. New York: Springer-Verlag.
- Draper, N., & Smith, N. (1981). Applied Regression Analysis (2nd ed.). New York: Wiley.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decisions under risk. Econometrica, 47, 263-291.
- Koopmans, T. C. (1960). Stationary ordinal utility and impatience. Econometrica, 28, 287-309.
- Loewenstein, G. (1987). Anticipation and the valuation of delayed consumption. The Economic Journal, 97, 666-684.
- Mellers, B. A. (1982). Equity judgment: A revision of Aristotelian views. Journal of Experimental Psychology: General, 111, 242-270.
- Meyers, R. F. (1976). Preferences over time. In R. L. Keeney & H. Raiffa (Eds.), Decisions with Multiple Objectives: Preferences and Value Tradeoffs. New York: Wiley.
- Miller, D. T., & Karniol, R. (1976). Coping strategies and attentional mechanisms in self-imposed and externally imposed delay situations. Journal of Personality and Social Psychology, 34, 310-316.

- Mischel, W., & Ebbeson, E. B. (1970). Attention in delay of gratification. Journal of Personality and Social Psychology, 16, 329-337.
- Mischel, W., & Grusec, J. (1967). Waiting for rewards and punishments: Effects of time and probability on choice. Journal of Personality and Social Psychology, 5, 24-31.
- Mischel, S., Grusec, J., & Masters, J. C. (1969). Effects of expected delay time on the subjective value of rewards and punishments. Journal of Personality and Social Psychology, 11, 363-373.
- Mischel, W., & Metzner, R. (1962). Preference for delayed reward as a function of age, intelligence and length of delay interval. Journal of Abnormal and Social Psychology, 64, 425-431.
- Miyamoto, J. M., & Eraker, S. A. (1988). A multiplicative model of the utility of survival duration and health quality. Journal of Experimental Psychology: General, 117, 3-20.
- Moore, B., Mischel, W., & Zeiss, A. (1976). Comparative effects of the reward stimulus and its cognitive representation in voluntary delay. Journal of Personality and Social Psychology, 34, 419-424.
- Nisan, M., & Minkowich, A. (1973). The effect of expected temporal distance on risk taking. Journal of Personality and Social Psychology, 25, 375-380.
- Renner, K. E. (1964). Delay of reinforcement: A historical overview. Psychological Bulletin, 61, 341-361.
- Rose, B. J., & Birnbaum, M. H. (1975). Judgments of differences and ratios of numerals. Perception & Psychophysics, 18, 194-200.
- Schoemaker, P. J. H. (1982). The expected utility model: Its variants, purposes, evidence and limitations. Journal of Economic Literature, 20, 529-563.
- Shanteau, J. C. (1975). An information integration analysis of risky decision making. In M. Kaplan & S. Swartz (Eds.), Human Judgment and Decision Processes. New York: Academic Press.
- Stevenson, M. K. (1986). A discounting model for decisions with delayed positive and negative outcomes. Journal of Experimental Psychology: General, 115, 131-154.
- Tarpy, R. M., & Sawabini, F. L. (1974). Reinforcement delay: A selective review of the last decade. Psychological Review, 81, 984-997.
- Tversky, A. (1969). Intransitivity of preferences. Psychological Review, 76, 31-48.

- von Winterfeldt, D., & Fischer, G. W. (1975). Multiattribute utility theory: Models and assessment procedures. In D. Wendt & C. Vlek (Eds.), Utility, Probability and Human Decision Making. Boston: D. Reidel Publishing.
- Winsberg, S., & Ramsay, J. O. (1981). Analysis of pairwise preference data using integrated B-splines. Psychometrika, 46, 171-186.
- Yates, B. T., & Mischel, W. (1979). Young children's preferred attentional strategies for delaying gratification. Journal of Personality and Social Psychology, 37, 286-300.
- Yates, J. F., & Watts, R. A. (1975). Preferences for deferred losses. Organizational Behavior and Human Performance, 13, 294-306.

Table 1

The Reliability of Responses for the Individual Subjects Across Four Sessions

Positive Outcome Task				
Subject	Risky	Rating	Preference	
	Variance	r	Variance	r
	Predicted	n=64	Predicted	n=64
07	95.76	.80	98.30	.92
10	94.74	.84	99.13	.93
13	96.12	.86	96.47	.90
27	94.26	.86	96.22	.89
56	96.87	.82	98.08	.93

Negative Outcome Task				
Subject	Risky	Rating	Preference	
	Variance	r	Variance	r
	Predicted	n=64	Predicted	n=64
07	95.39	.78	98.26	.91
10	90.58	.68	99.14	.88
13	96.81	.86	99.02	.95
27	96.26	.88	98.14	.94
56	95.53	.79	98.14	.90

Table 2

The Subjective Values Obtained for the Individual Subjects for the
Magnitude and Delay of the Outcomes

Positive Outcome Task							
		Percent Funding			Time in Years		
Subj			Risky	Riskfree		Risky	Riskfree
07	1	25%	1.28	1.28	6 mos.	.26	.26
	2	50%	3.16	5.48	1 yr.	.42	.65
	3	75%	7.41	8.32	2 yrs.	1.01	1.21
	4	100%	11.68	11.68	4 yrs.	1.83	1.83
10	1	25%	2.36	2.36	6 mos.	.28	.28
	2	50%	3.94	7.43	1 yr.	.44	.50
	3	75%	8.25	9.72	2 yrs.	1.08	.94
	4	100%	11.16	11.16	4 yrs.	1.34	1.34
13	1	25%	2.76	2.76	6 mos.	.015	.015
	2	50%	4.80	6.62	1 yr.	.023	.024
	3	75%	7.56	10.32	2 yrs.	.048	.040
	4	100%	11.02	11.02	4 yrs.	.055	.055
27	1	25%	.14	.14	6 mos.	.02	.02
	2	50%	.27	4.79	1 yr.	.03	.03
	3	75%	4.37	9.99	2 yrs.	.04	.05
	4	100%	14.14	14.14	4 yrs.	.06	.06

(table continues)

Positive Outcome Task

Subj	Percent Funding				Time in Years	
			Risky	Riskfree	Risky	Riskfree
56	1	25%	1.75	1.75	6 mos.	.11
	2	50%	4.33	5.74	1 yr.	.16
	3	75%	8.78	9.58	2 yrs.	.28
	4	100%	10.44	10.44	4 yrs.	.39

Negative Outcome Task

Subj	Pay Periods				Time in Years	
			Risky	Riskfree	Risky	Riskfree
07	1	1 yr.	4.27	4.27	2 yrs.	.60
	2	3 yrs.	6.45	6.62	4 yrs.	1.10
	3	4 yrs.	8.73	7.87	6 yrs.	1.33
	4	6 yrs.	8.75	8.75	8 yrs.	2.66
10	1	1 yr.	2.97	2.97	2 yrs.	.24
	2	3 yrs.	4.27	5.26	4 yrs.	.31
	3	4 yrs.	8.33	8.73	6 yrs.	.40
	4	6 yrs.	10.57	10.57	8 yrs.	.58
13	1	1 yr.	4.83	4.83	2 yrs.	.62
	2	3 yrs.	5.98	6.96	4 yrs.	1.00
	3	4 yrs.	7.74	8.35	6 yrs.	1.21
	4	6 yrs.	9.21	9.21	8 yrs.	1.81

(table continues)

Negative Outcome Task

			Pay Periods		Time in Years		
Subj			Risky	Riskfree		Risky	Riskfree
27	1	1 yr.	3.56	3.56	2 yrs.	.23	.23
	2	3 yrs.	5.87	6.49	4 yrs.	.31	.30
	3	4 yrs.	8.37	7.90	6 yrs.	.34	.33
	4	6 yrs.	9.62	9.62	8 yrs.	.51	.51
56	1	1 yr.	2.99	2.99	2 yrs.	.46	.46
	2	3 yrs.	5.74	6.70	4 yrs.	.62	.73
	3	4 yrs.	8.12	8.99	6 yrs.	.73	.80
	4	6 yrs.	9.88	9.88	8 yrs.	1.18	1.18

Note: Time Scale Value is the reciprocal of the discounting weight.

The Sample of Funding Program Attributes

Percentage of Support

x - This level of support was available.

Example: See Figure 2

All support combinations for Years 1, 2, and 3 were used.

Example:	Year 1	Year 2	Year 3
	10%	90%	40%

All two way combinations were evaluated.

Example:	Year 3	Year 4
	90%	10%

Example: Year 2

67

Table 4

The Discounting Weights Estimated for the Groups and Individuals

Averaging Model						
Group R (N=10)		Group L (N=13)		Individual		
95.33%		97.20%		Subject 11		
$w_1 = .76$		$w_1 = .40$		$w_1 = .27$		
$w_2 = .66$		$w_2 = .43$		$w_2 = .27$		
$w_3 = .67$		$w_3 = .45$		$w_3 = .26$		
$w_4 = .67$		$w_4 = .45$		$w_4 = .22$		
				$w_0 = 1$		
Adding Model						
Group R (N=13)			Group L (N=9)			
92.06%			92.53%			
$w_1 = .07$			$w_1 = .10$			
$w_2 = .07$			$w_2 = .09$			
$w_3 = .07$			$w_3 = .10$			
$w_4 = .06$			$w_4 = .10$			
Individuals						
	Subj 10	Subj 26	Subj 57	Subj 61	Subj 62	Subj 65
w_1	.10	.10	.10	.07	.07	.07
w_2	.08	.06	.11	.07	.07	.07
w_3	.11	.05	.09	.07	.07	.08
w_4	.15	.07	.08	.07	.07	.07

Table 5

Reliability and Discounting Weights for the Individual Subjects

	Subject (Correlation)			Discounting Weights per Year			
	N=187						
	Sub11	Sub10	Sub26	Sub57	Sub61	Sub62	Sub65
	(.91)	(.89)	(.94)	(.89)	(.94)	(.96)	(.93)
Time							
1	.27	.10	.07	.10	.07	.07	.07
2	.24	.08	.05	.11	.07	.07	.07
3	.26	.11	.06	.09	.07	.07	.08
4	.22	.15	.10	.08	.07	.07	.07
Events							
	Set Size Weights						
1	.73	.10	.10	.10	.10	.10	.10
2	.51	.14	.09	.33	.22	.10	.29
3	.38	.27	.08	.76	.56	.10	.33
4	.29	.42	.08	1.00	1.79	.10	.38

Figure Captions

Figure 1. a) Conceptual model for the evaluation of risky funding programs.
b) Conceptual model for the comparison of riskfree funding programs.

Figure 2. The fit of the multiplicative model for the risky funding program described as positive evaluations.

Figure 3. The mean preference ratings obtained for the funding programs.

Figure 4. The form of the response function describing strength of preference for positive outcomes.

Figure 5. The scale values for the funding programs estimated for the preference task plotted as a function of their attributes: support and time.

Figure 6. Psychophysical functions for the amount of funding and work time requirement.

Figure 7. The fit of the multiplicative model for the risky funding program described as negative evaluations.

Figure 8. The form of the response function describing strength of preference for negative outcomes.

Figure 9. The scale values for the loan programs estimated from the preference task plotted as a function of their attributes: pay-back period and loan time.

Figure 10. Psychophysical functions for the pay-back periods and loan time.

Figure 11. An example of a multiple outcome work-study program.

Figure 12. The response function for a) Group R and b) Group L describing the fit of the averaging model.

Figure 13. The observed judgments for work-study programs offering support for two years and single years for Group R (a & b) Adding Strategy (top panels c & d) Averaging Strategy (bottom panels).

Figure 14. The observed judgments for work study programs offering support for two years and single years for Group L (a & b) Adding Strategy (top panel c & d) Averaging Strategy (bottom panel).

Figure 15. Set size variations in the observed ratings for the group classified as average a) in Group R b) in Group L.

Figure 16. The response function for a) Group R and b) Group L describing the fit of the set size model.

Figure 17. The response function for a) Group R and b) Group L describing the fit of the additive model.

Figure 18. Set size variations in the observed ratings for the group classified as additive a) in Group R and b) in Group L.

Figure 19. The response function for a) Group R and b) Group L describing the fit of the set size model.

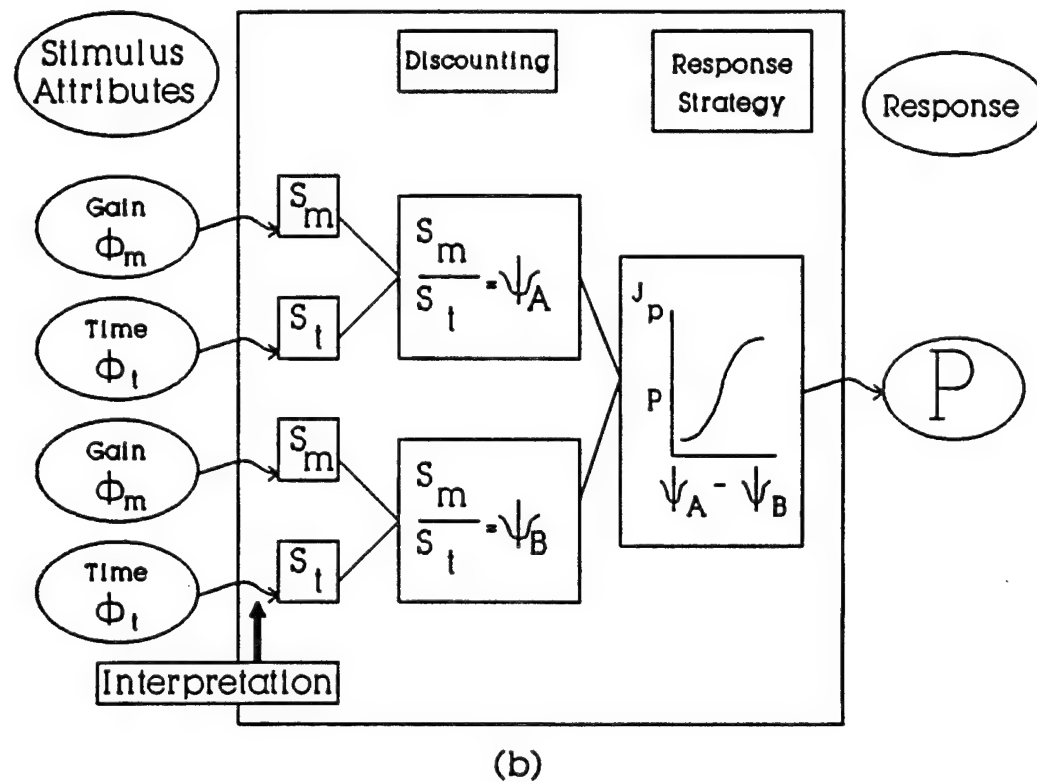
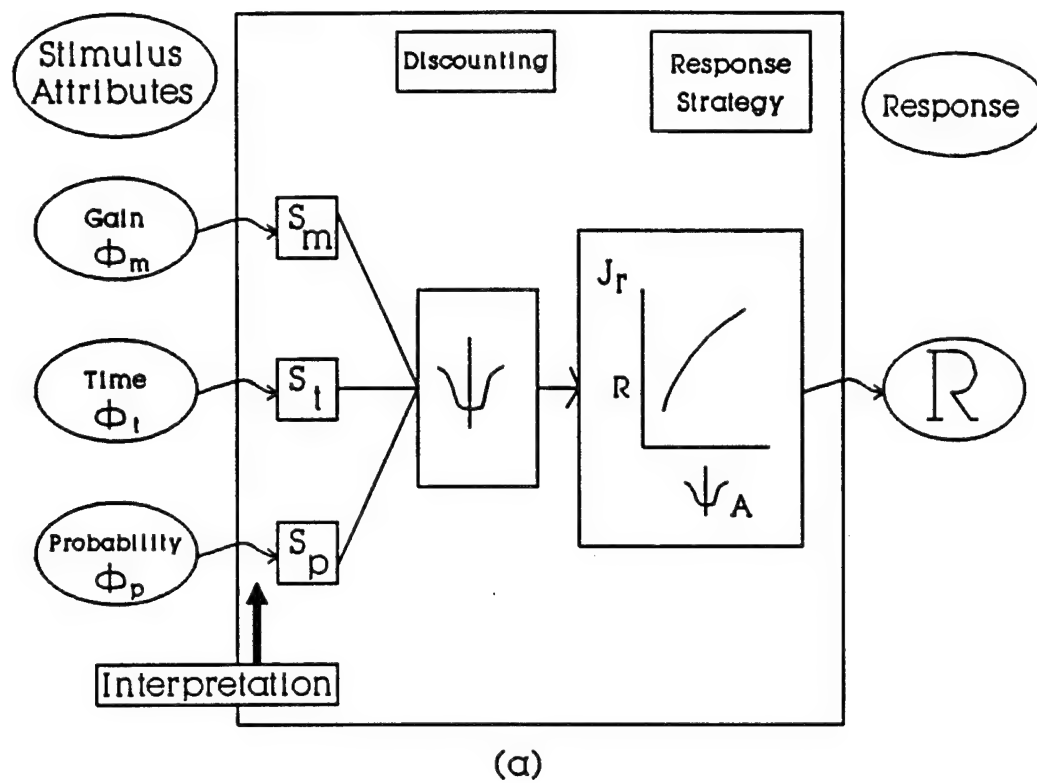


Figure 1

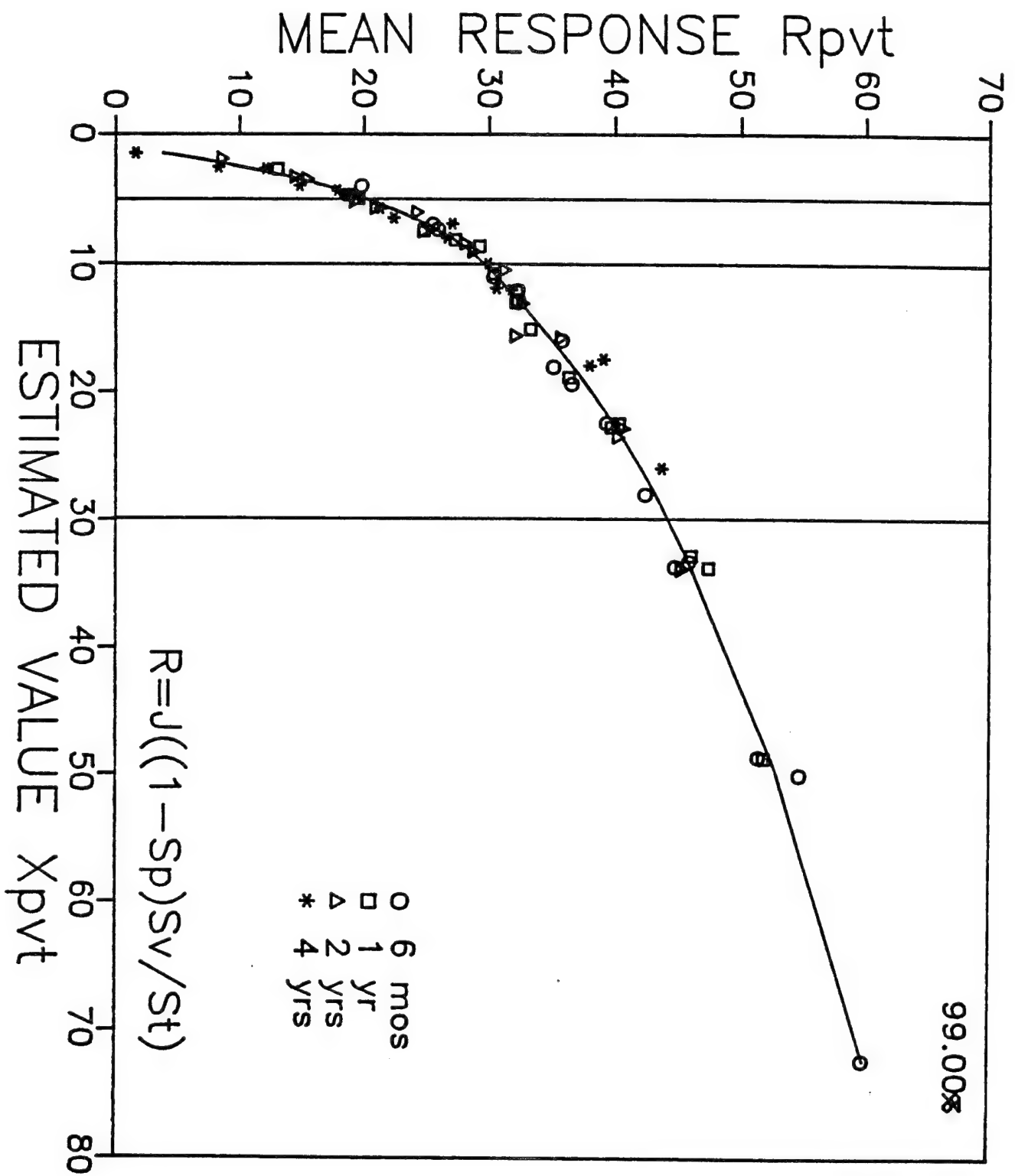
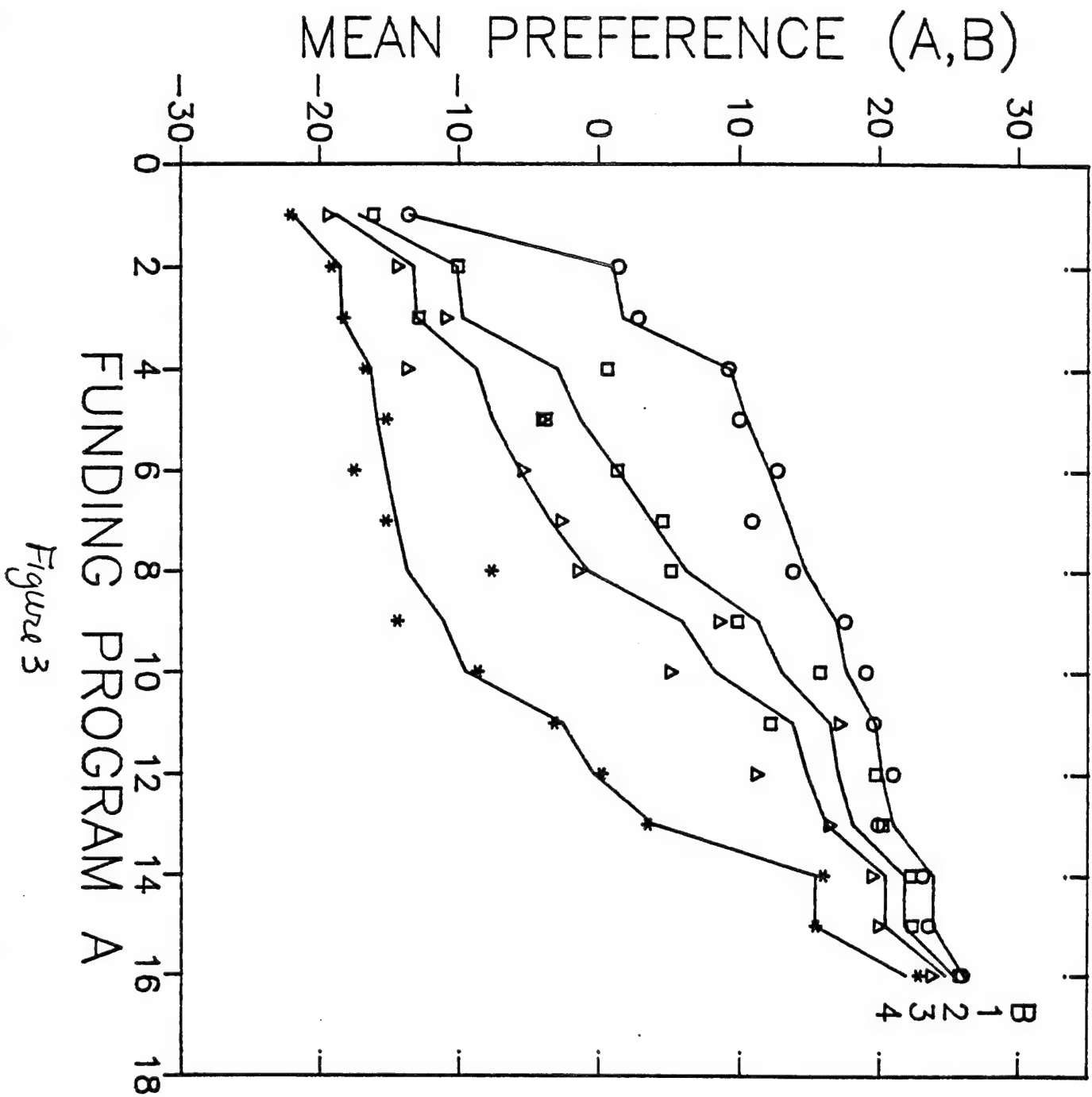


Figure 2



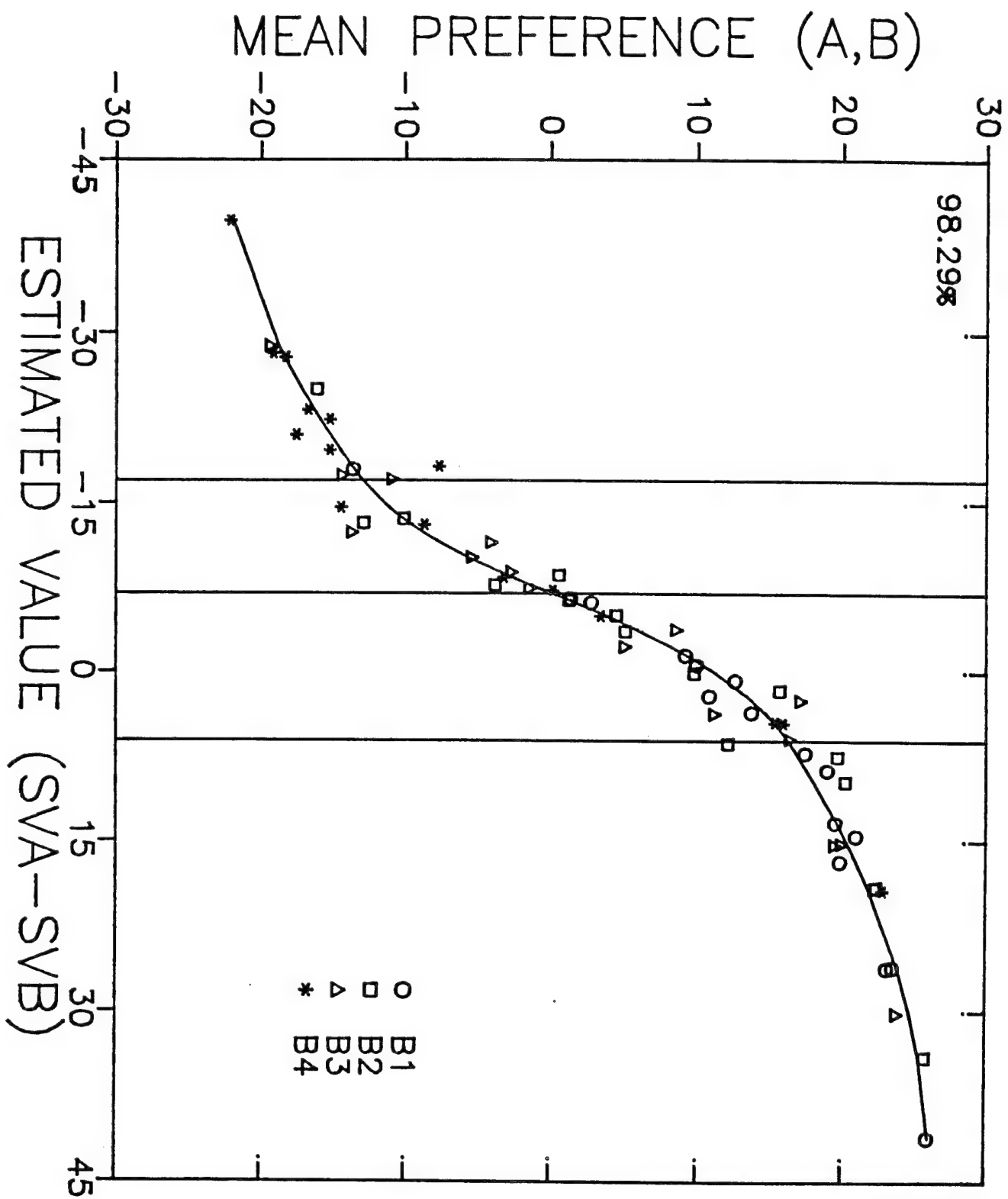


Figure 4

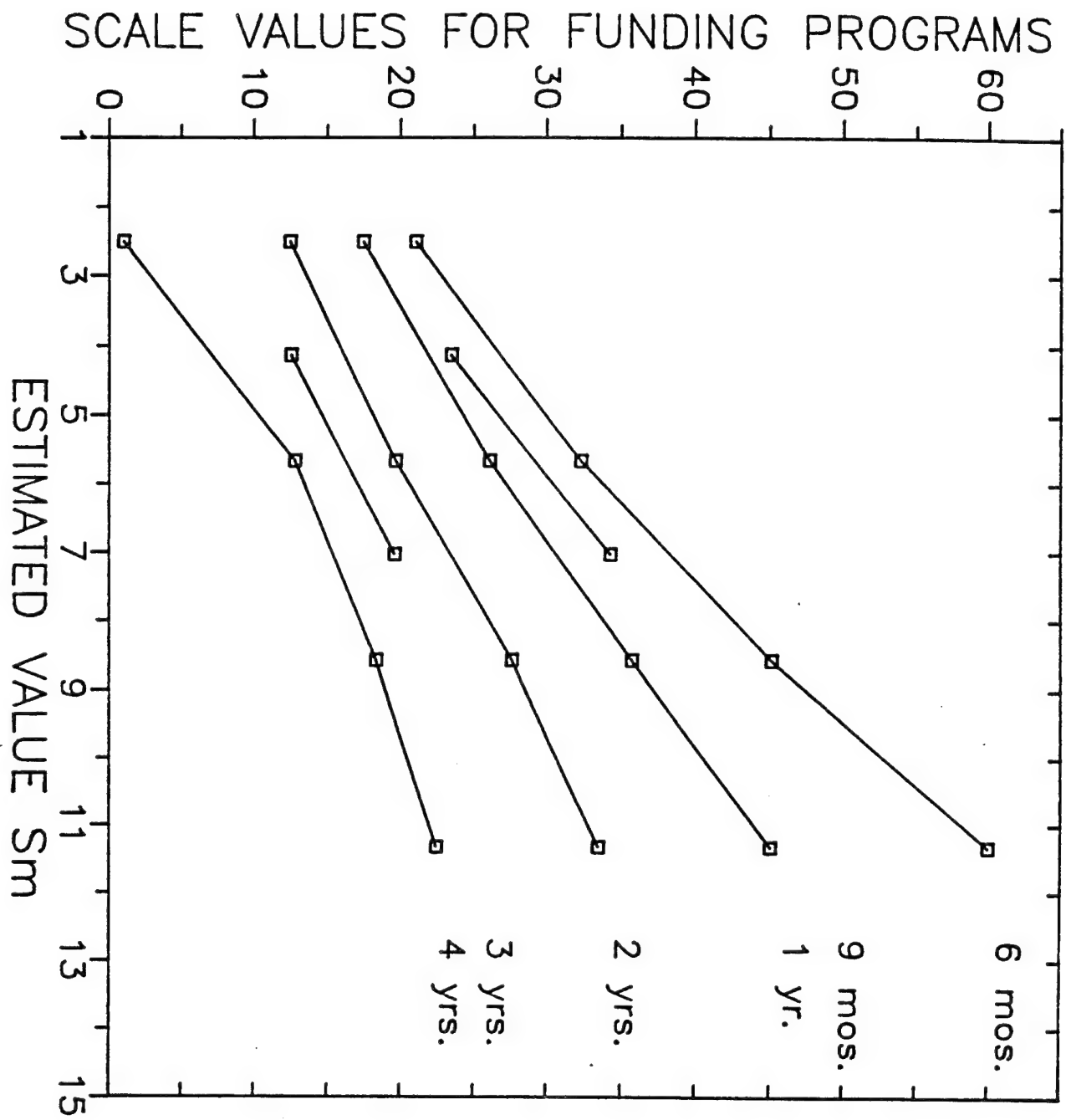


Figure 5

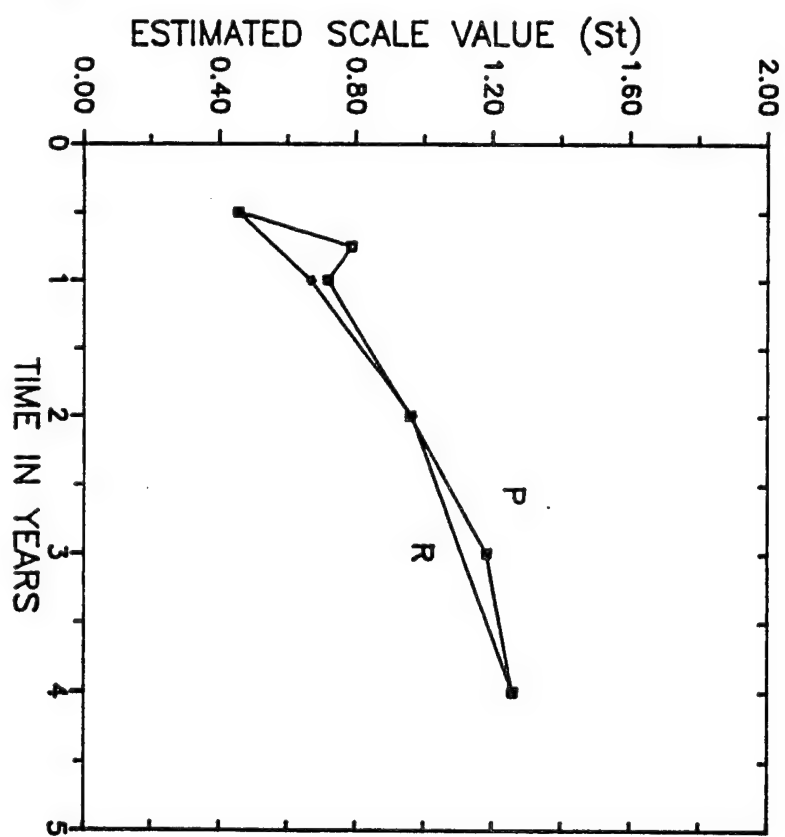
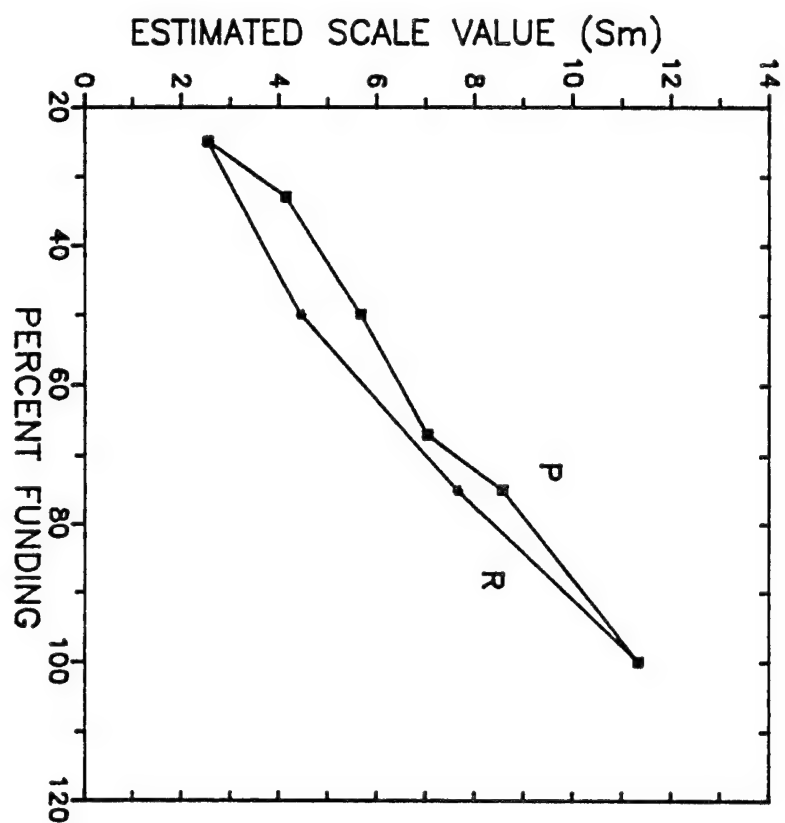


Figure 6

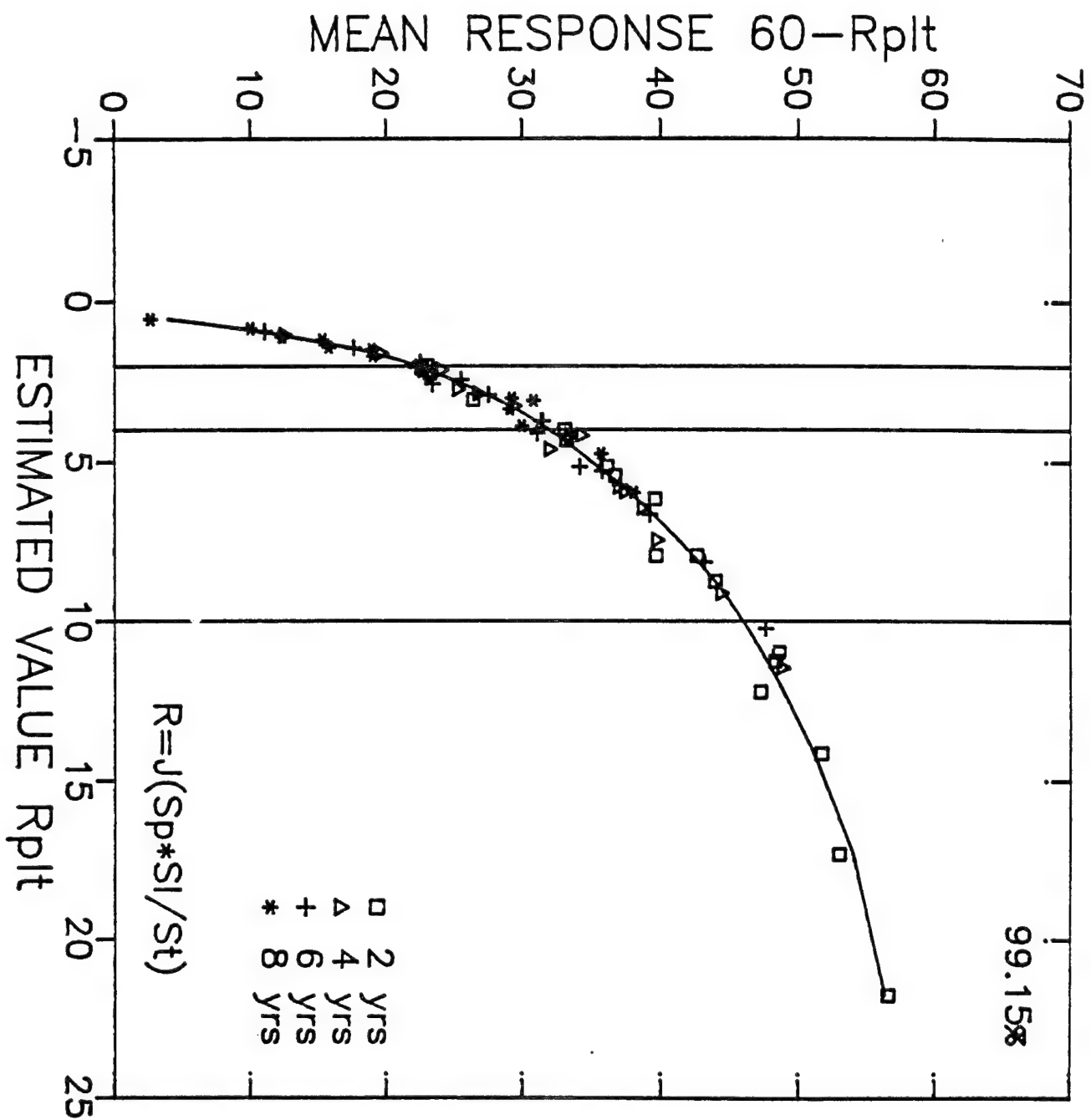


Figure 7

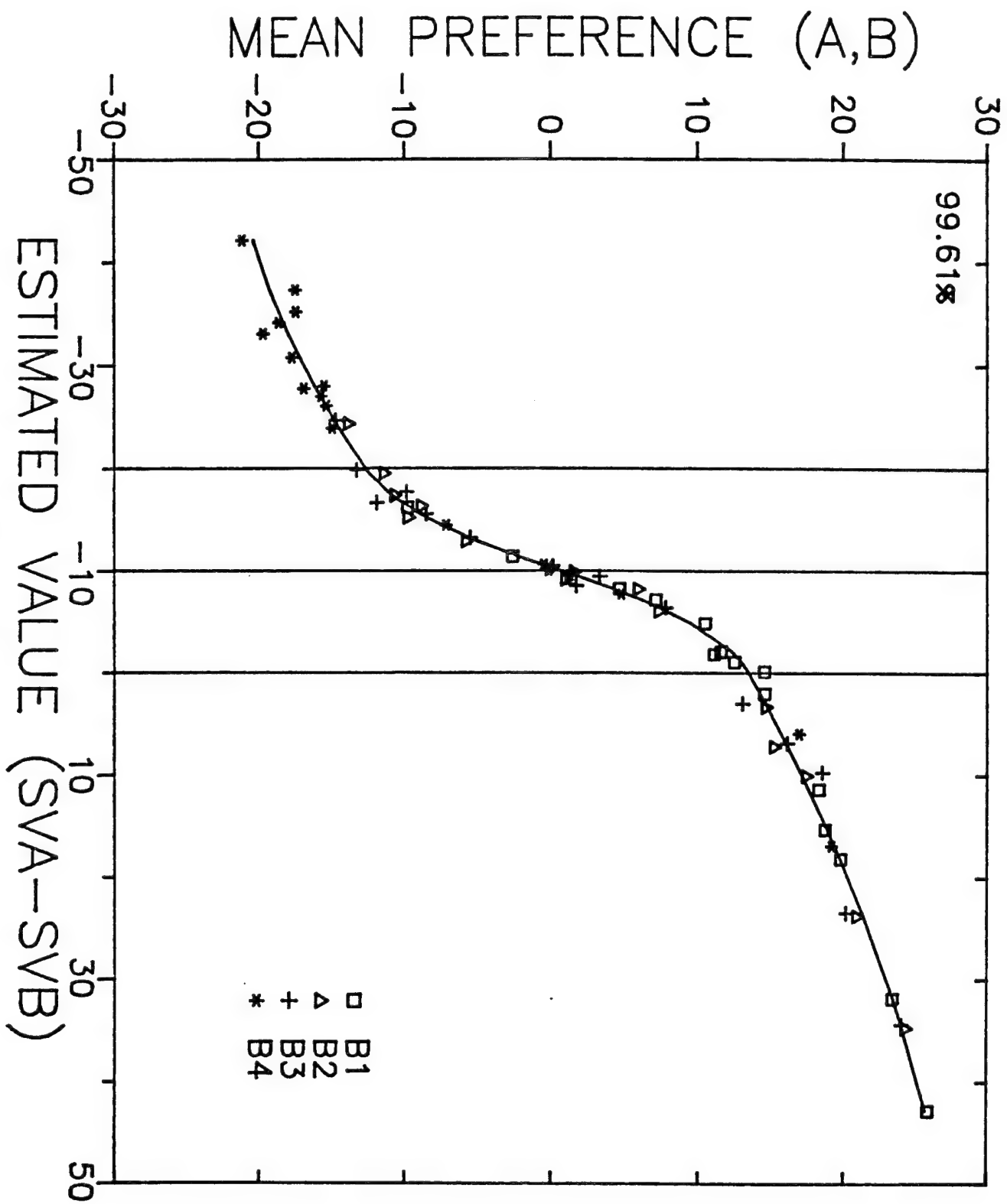


Figure 8

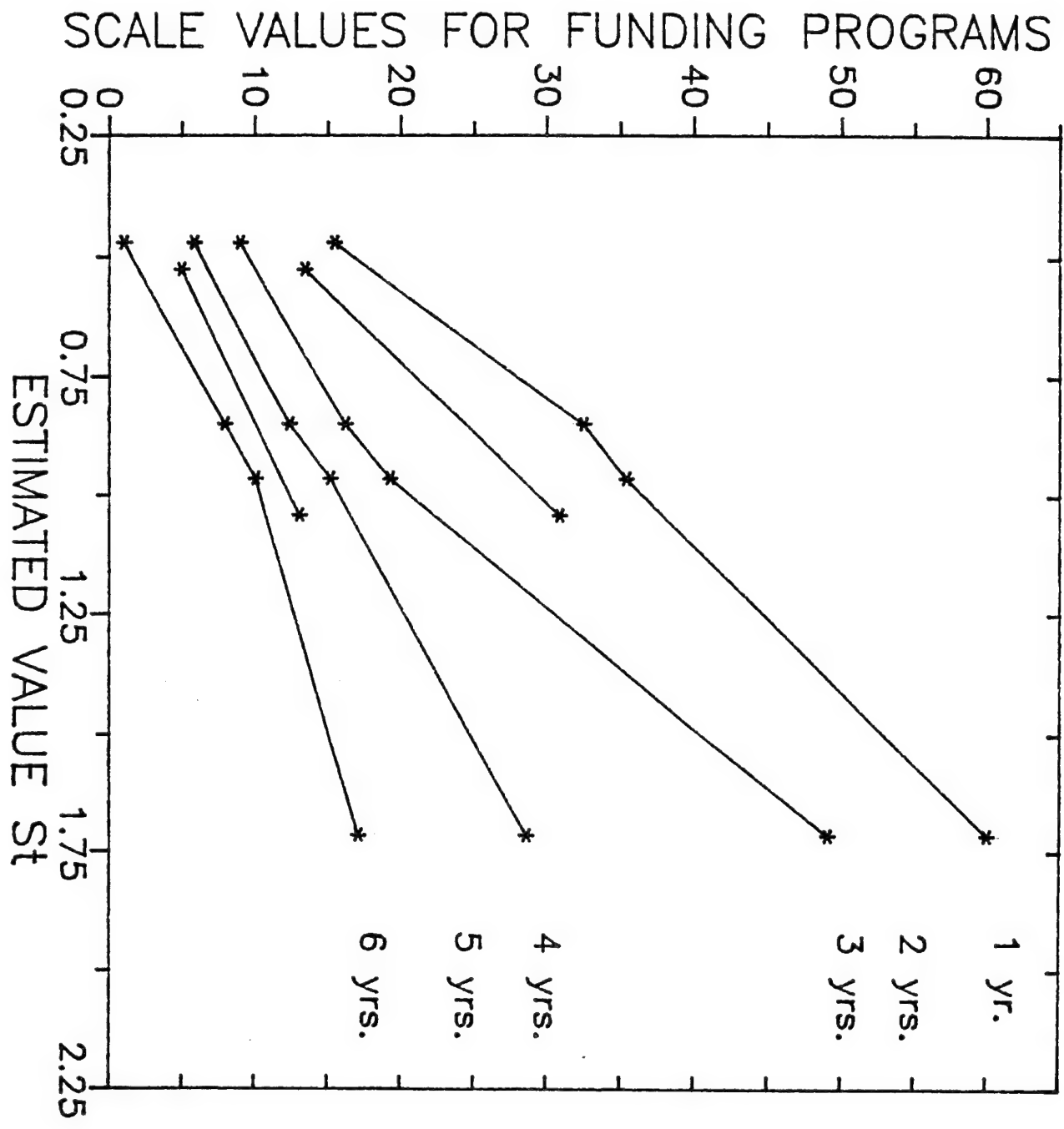


Figure 9

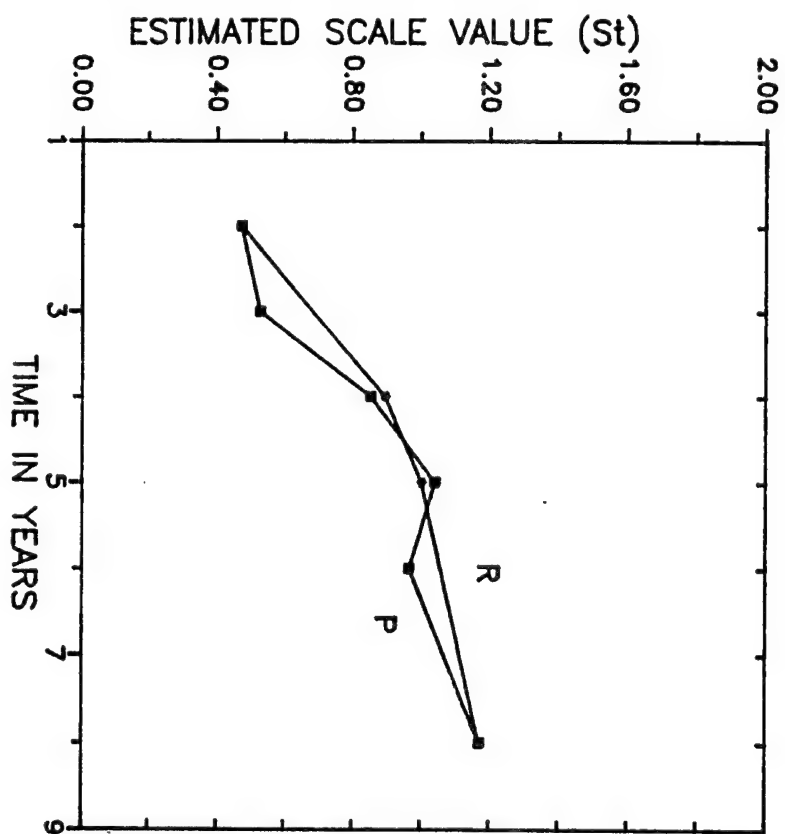
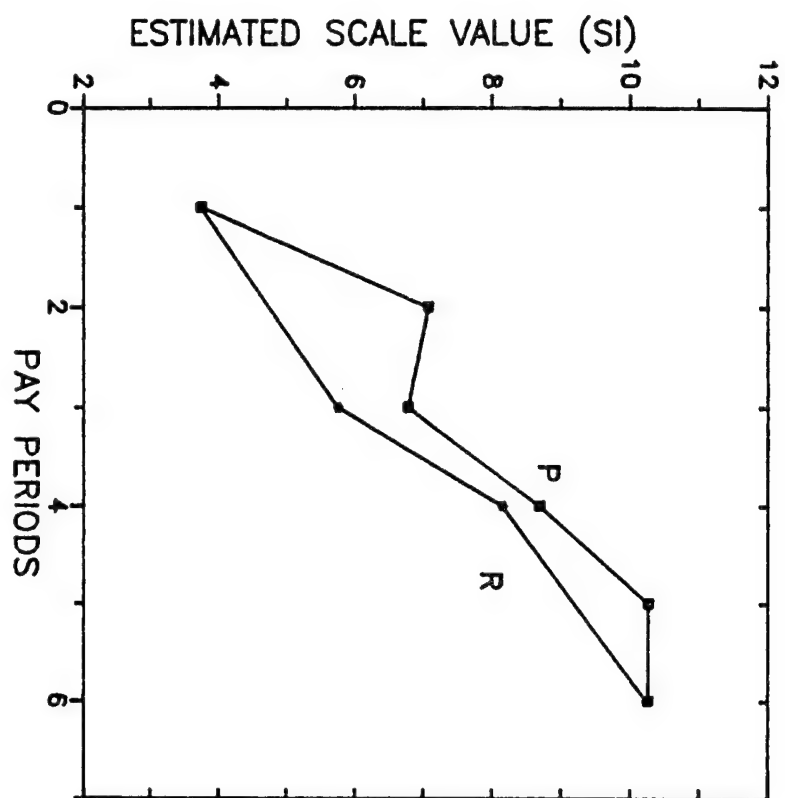


Figure 10

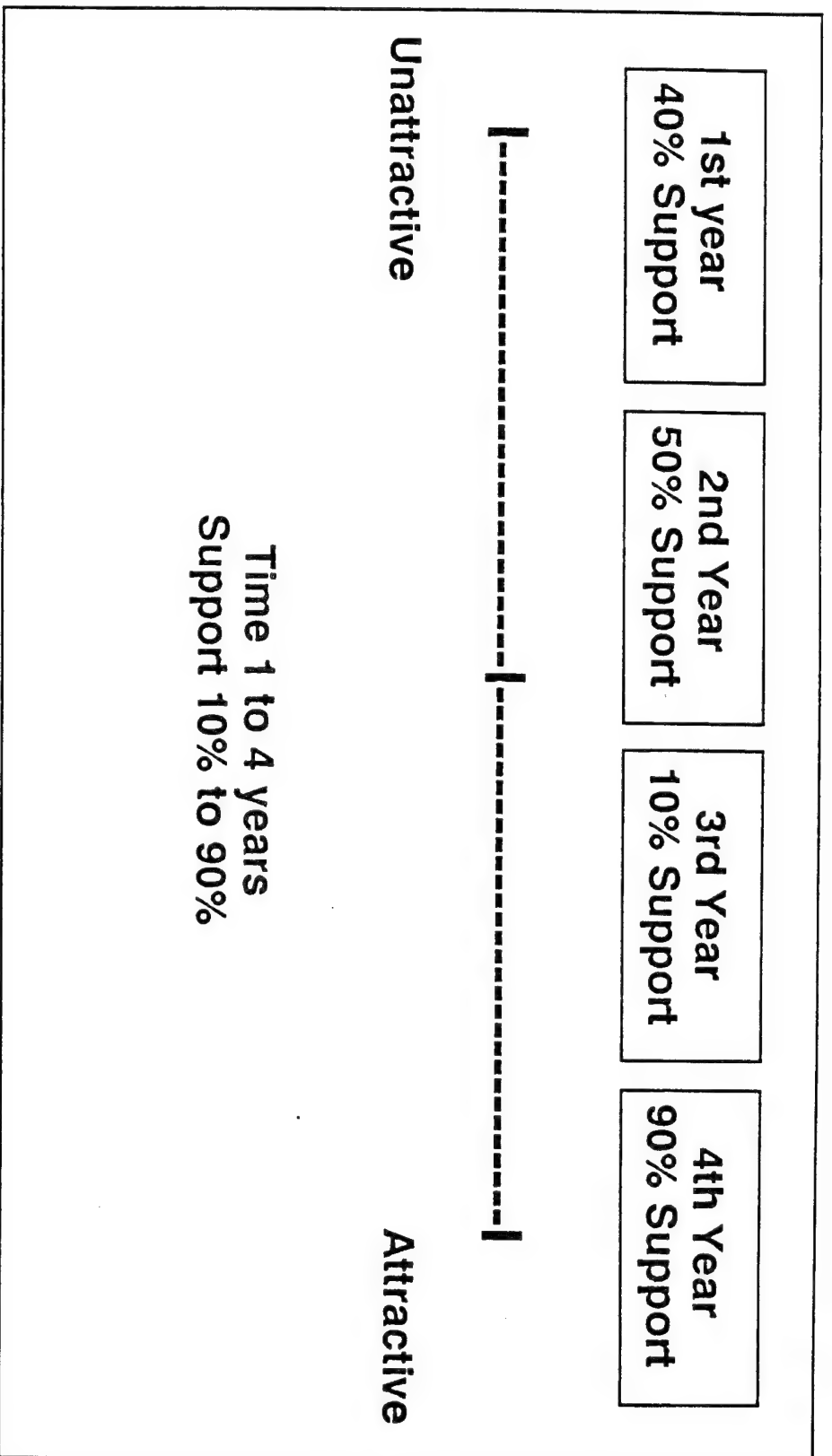


Figure 11

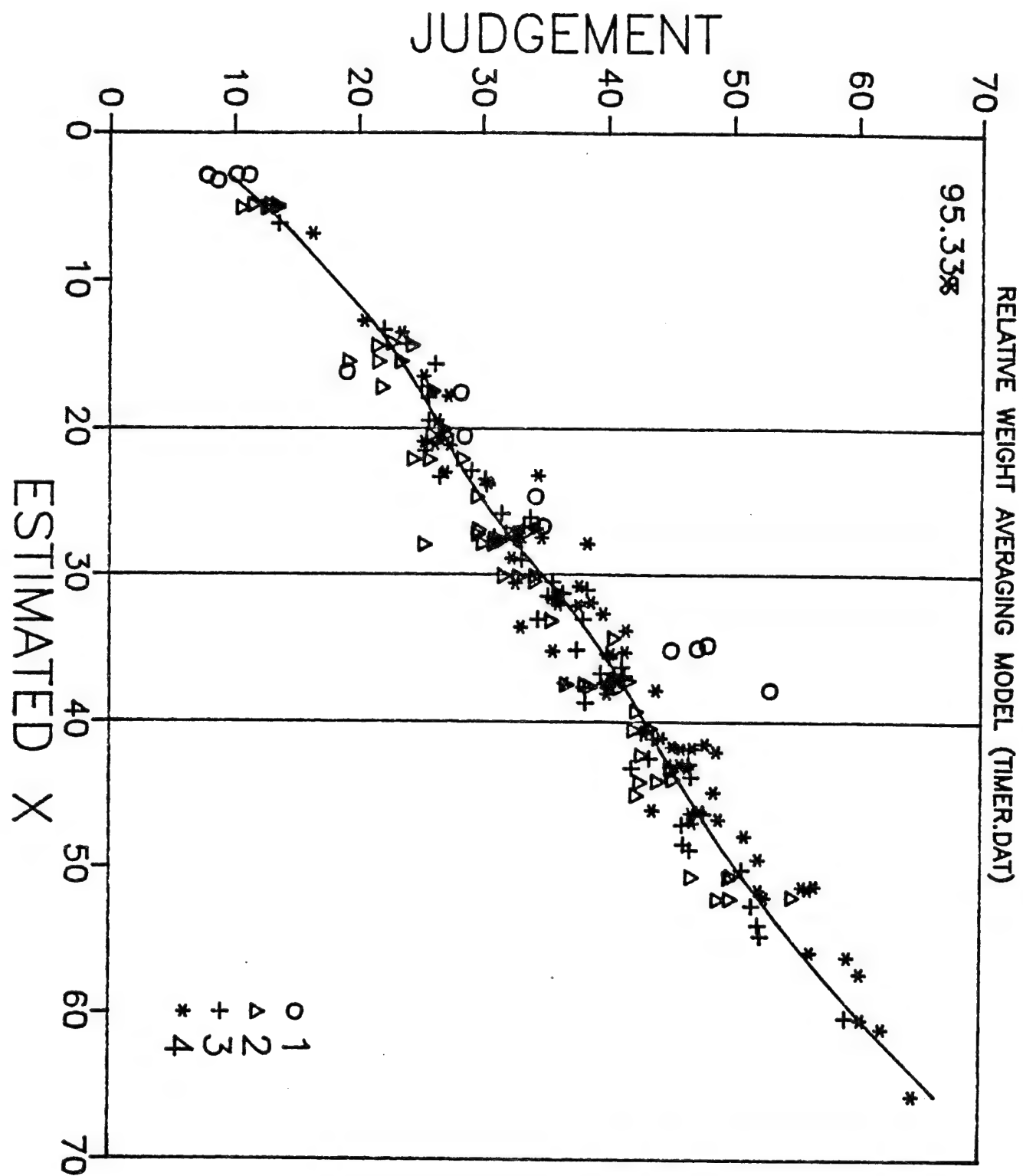


Figure 12a

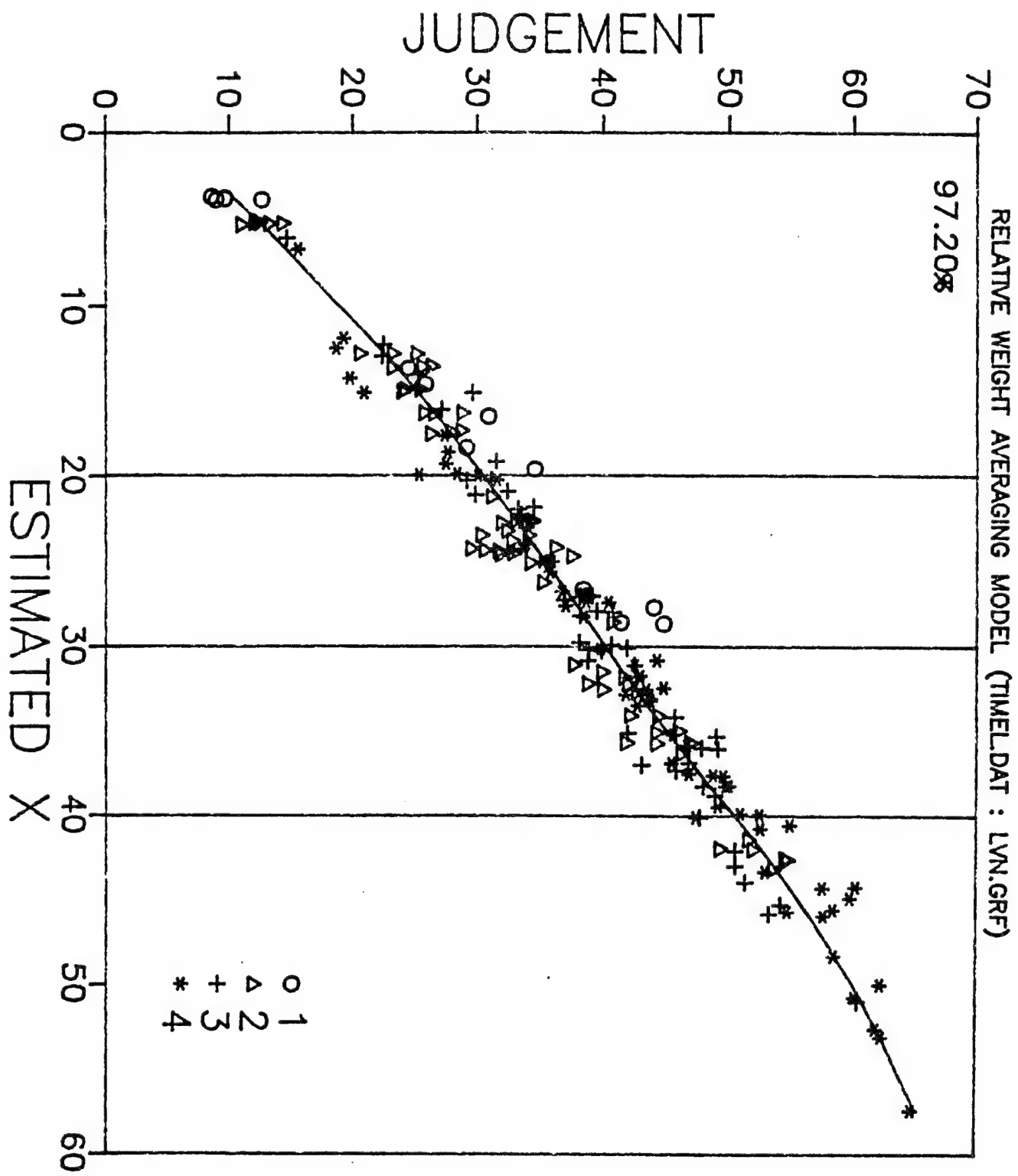


Figure 12b

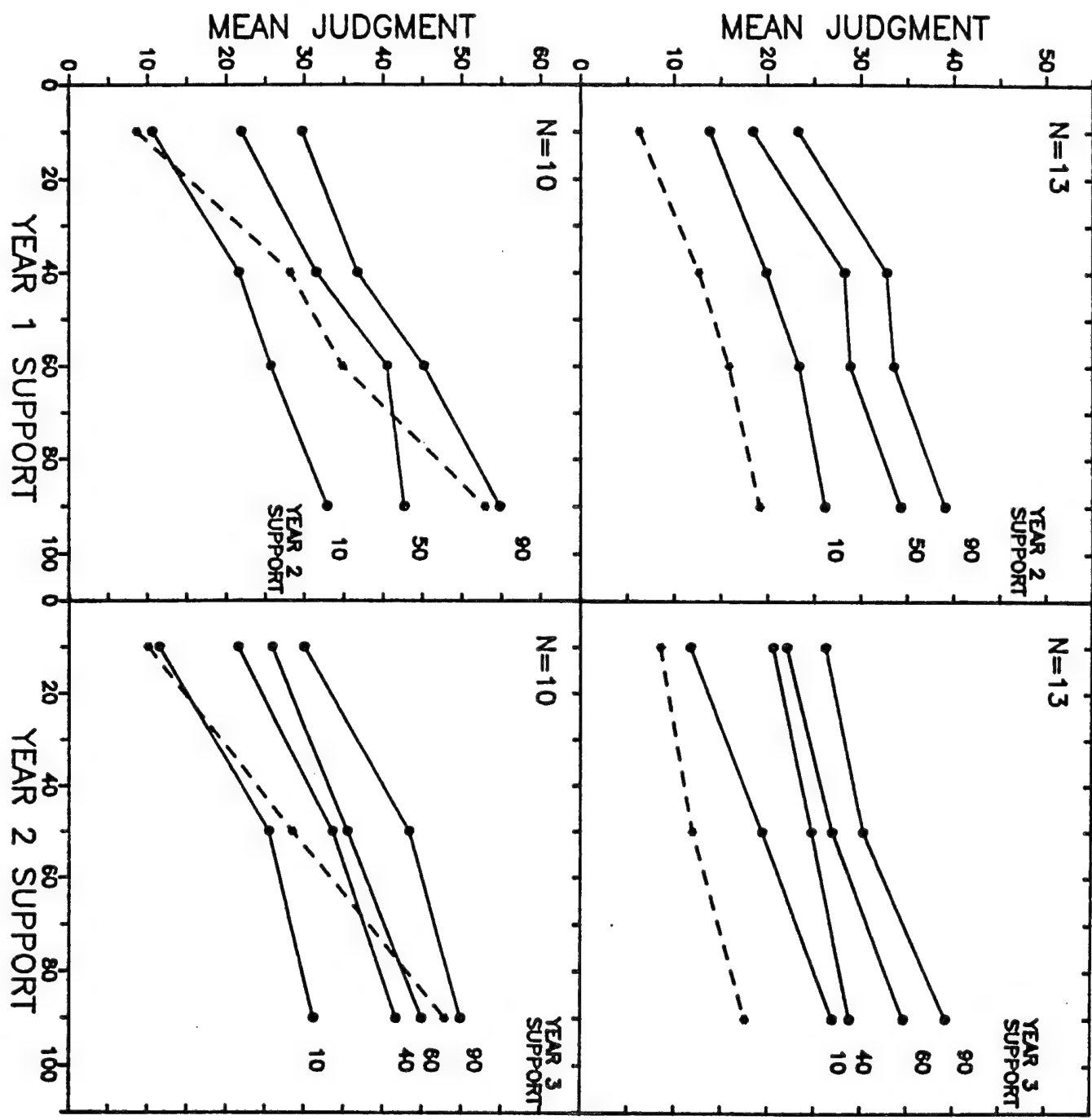


Figure 13

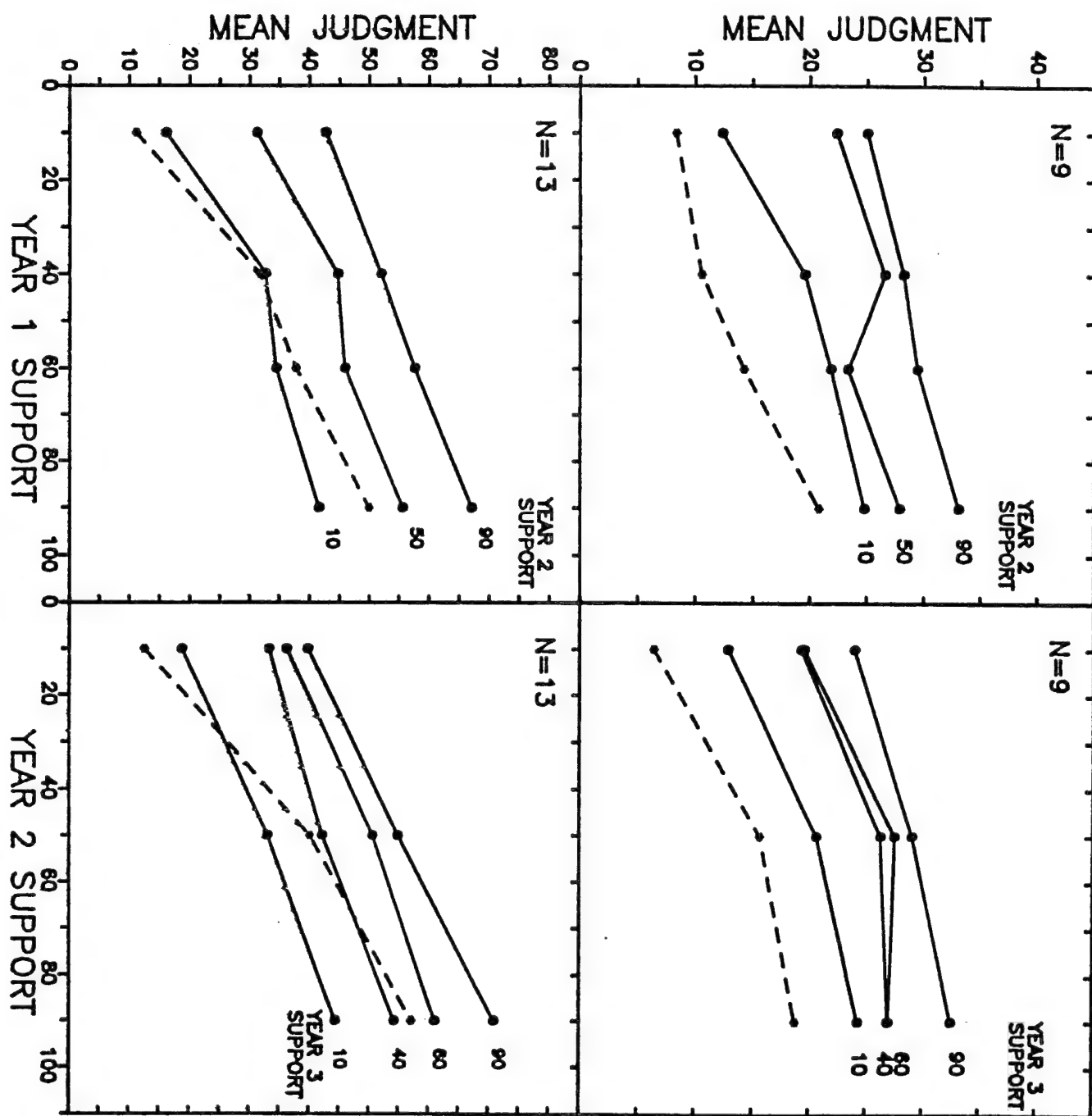


Figure 14

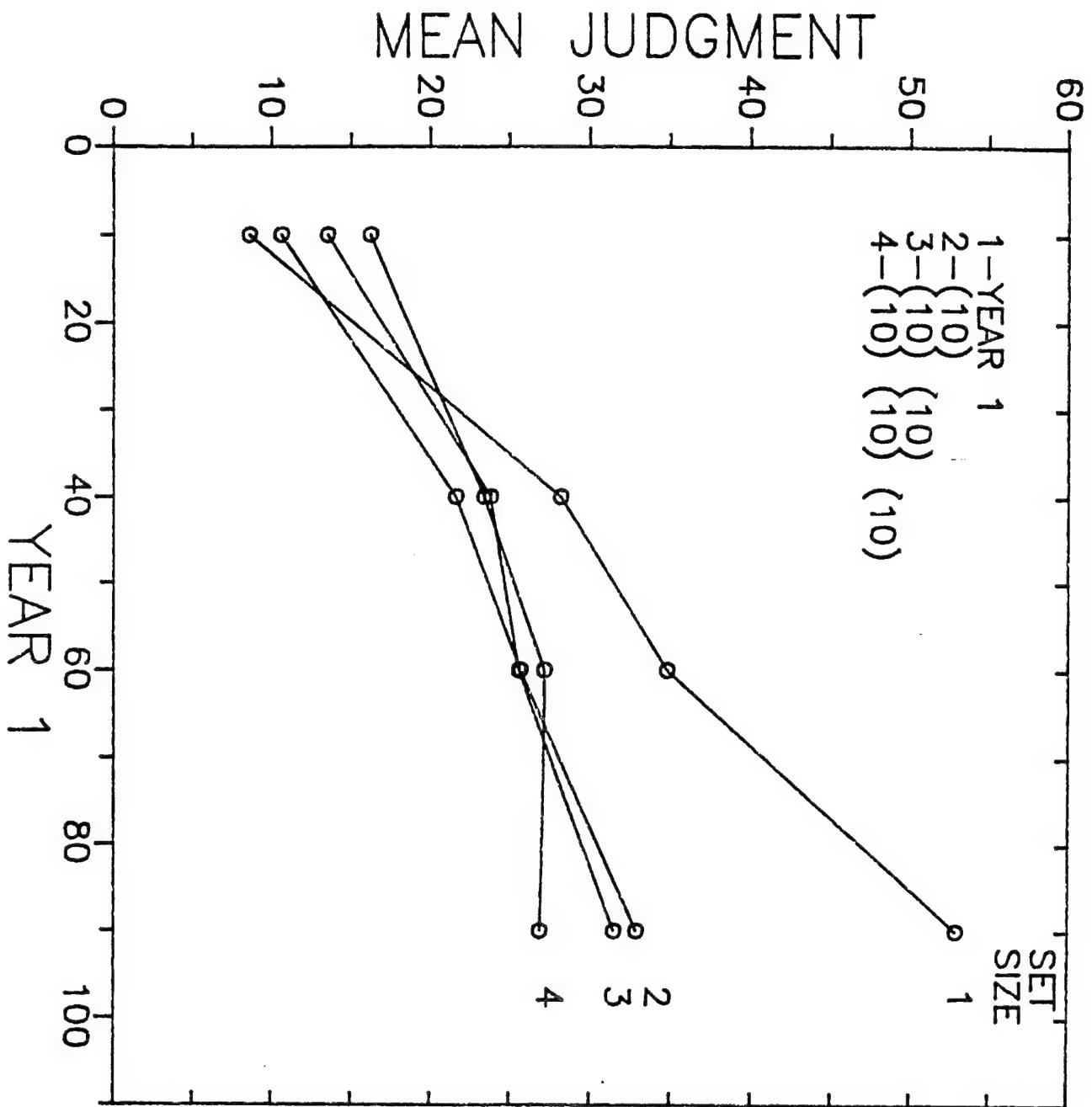
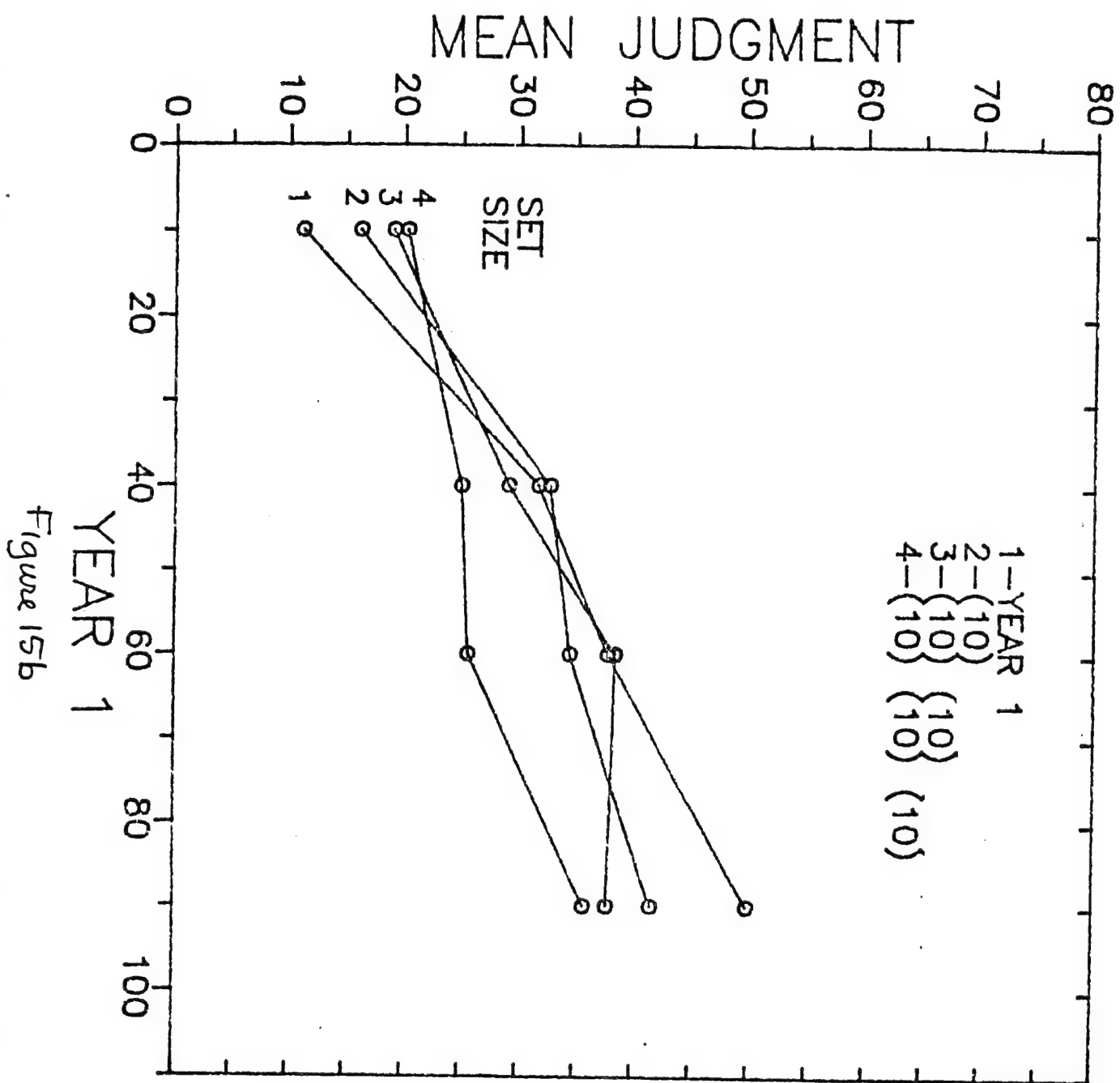


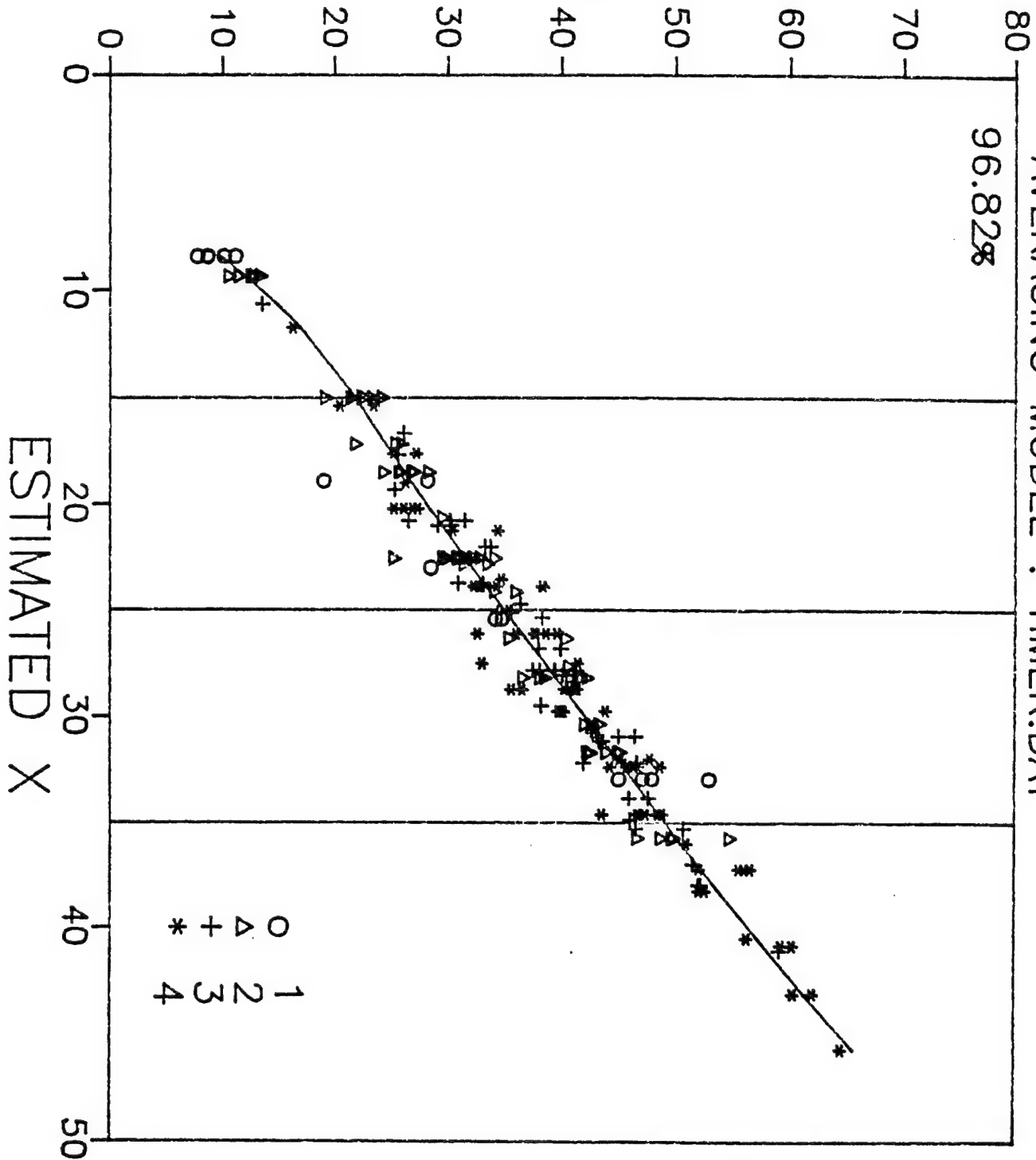
Figure 15a



AVERAGING MODEL : TIMER.DAT

96.82%

JUDGMENT



ESTIMATED X
Figure 16a

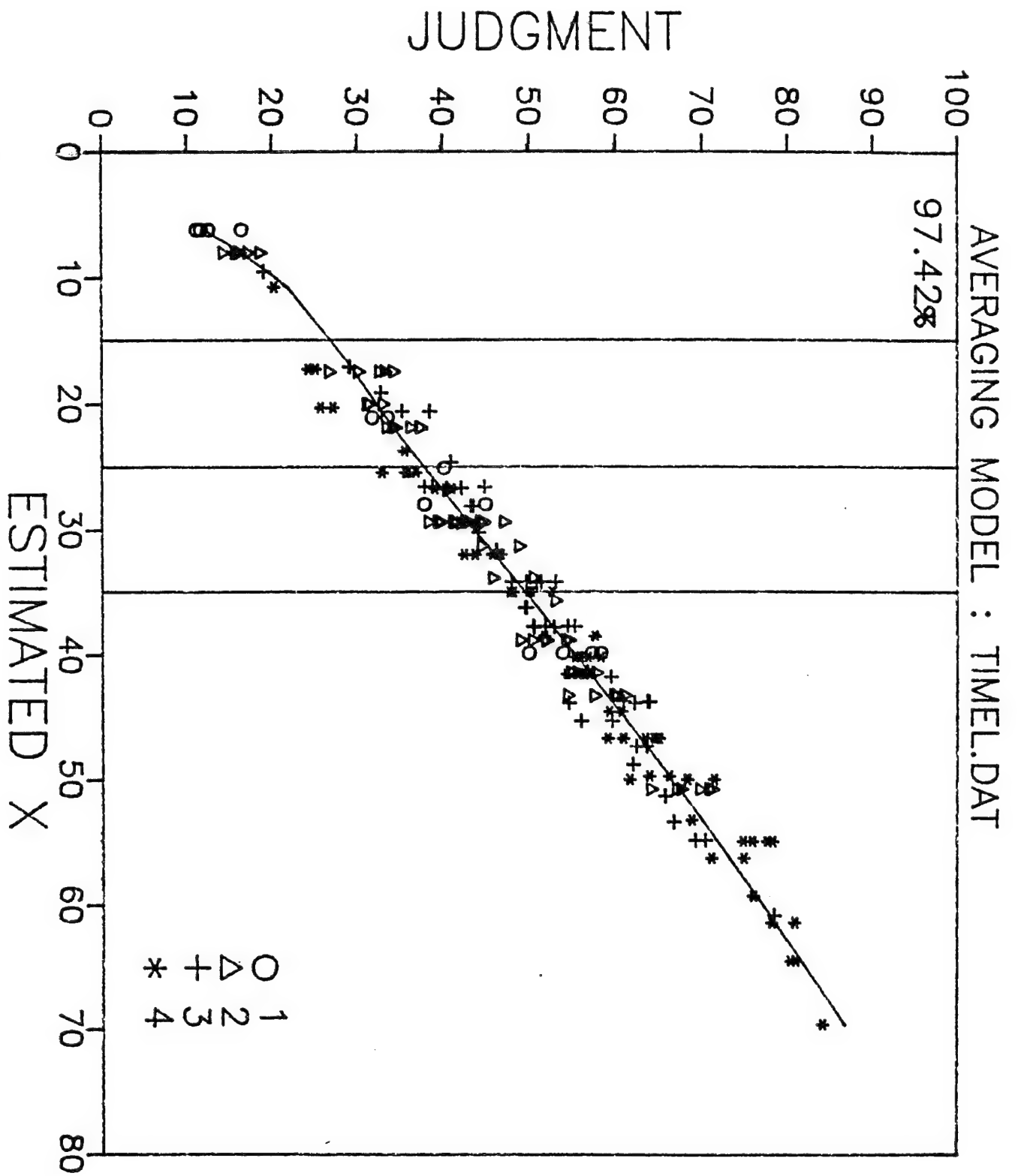


Figure 166

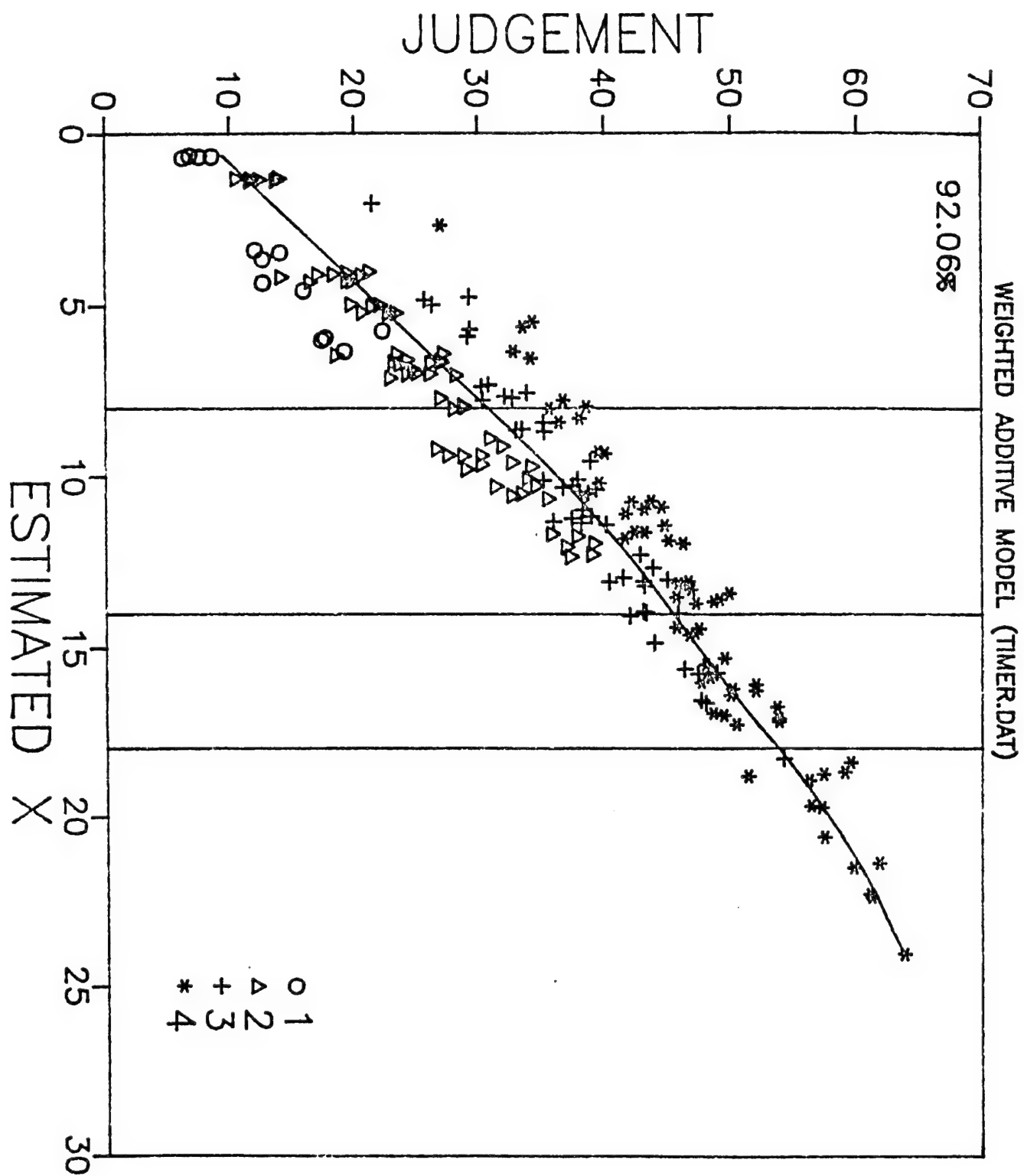


Figure 17a

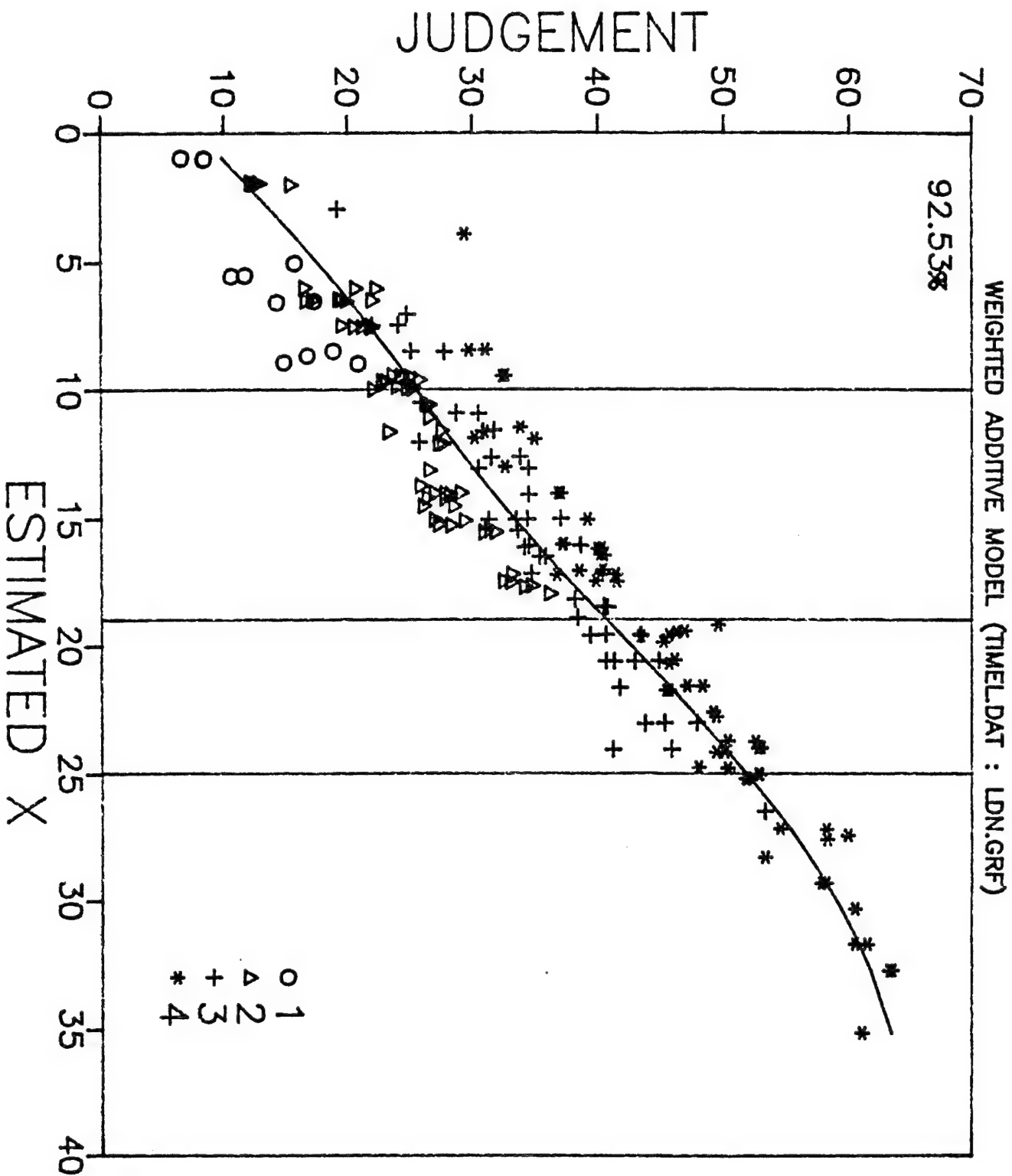


Figure 17b

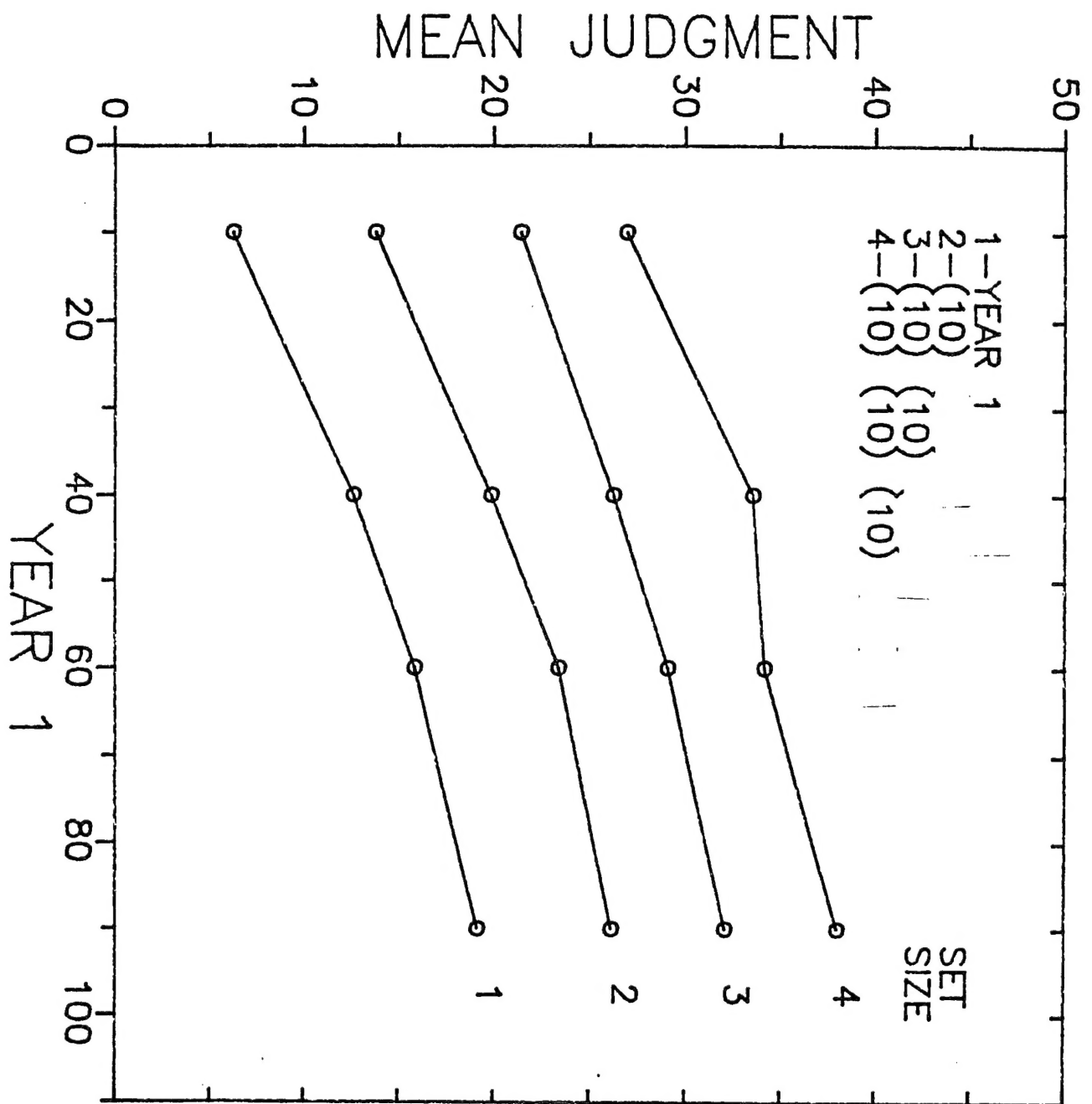


Figure 18a

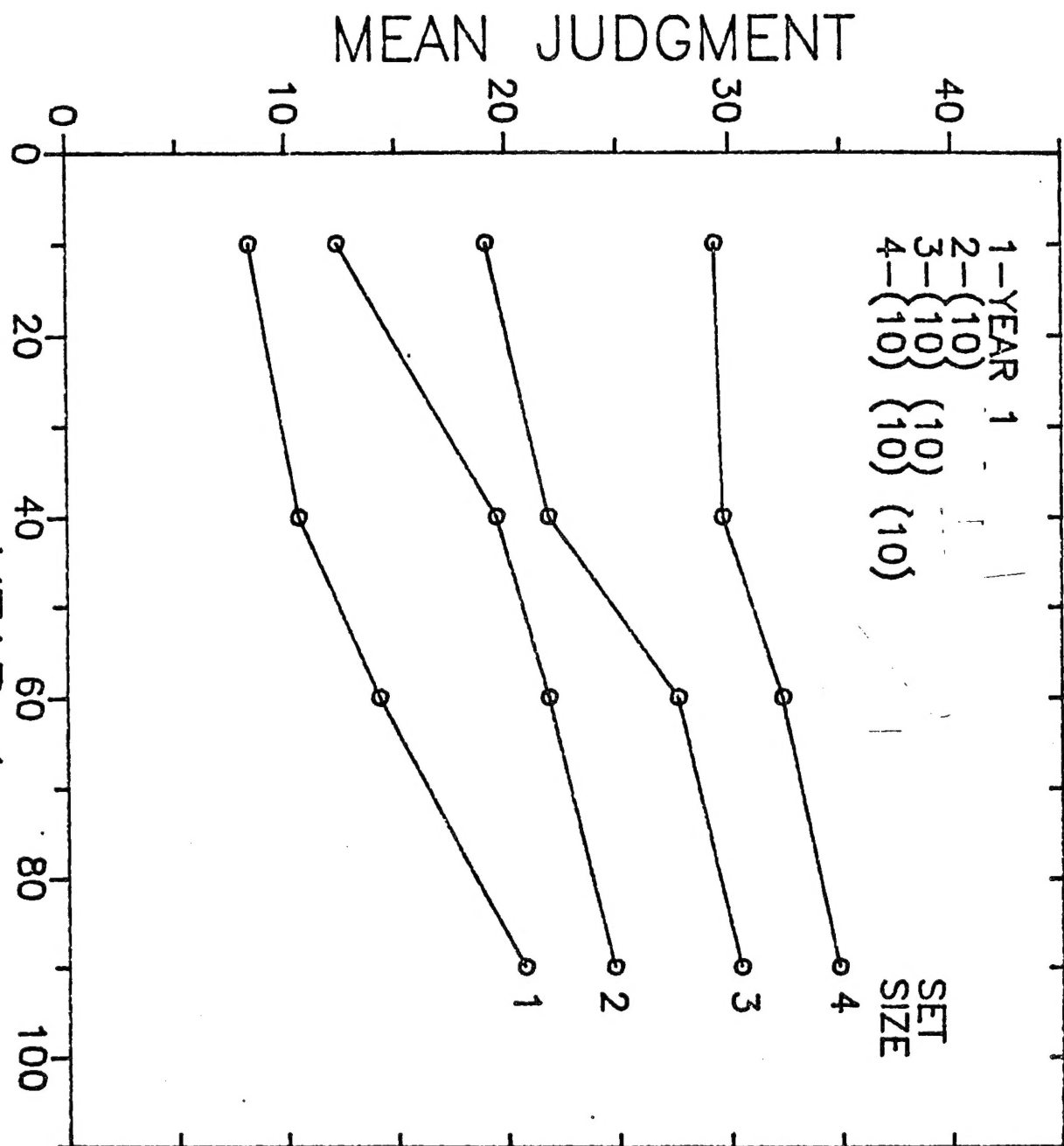


Figure 18b

ADDING MODEL : TIMER.DAT

98.15%

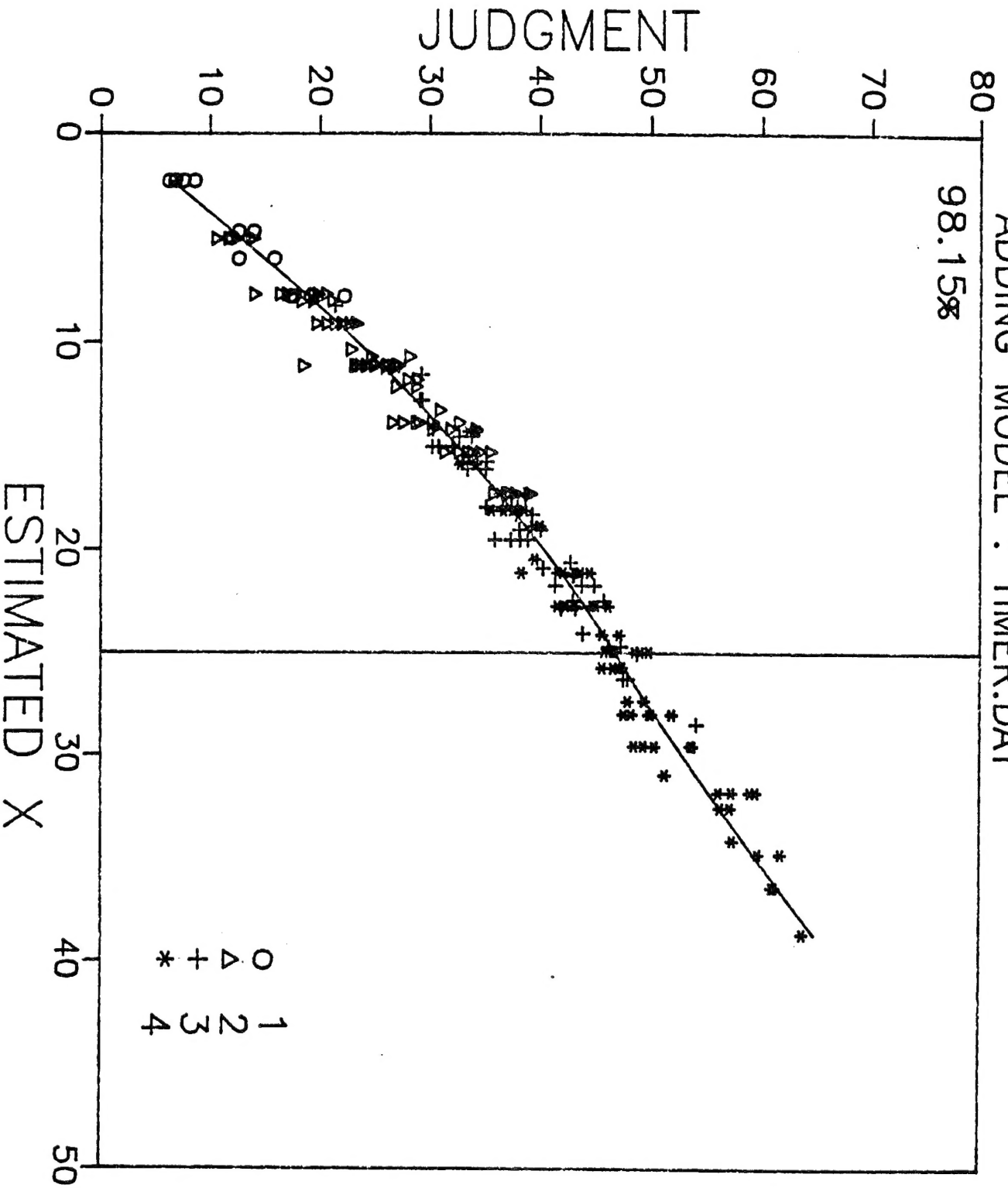
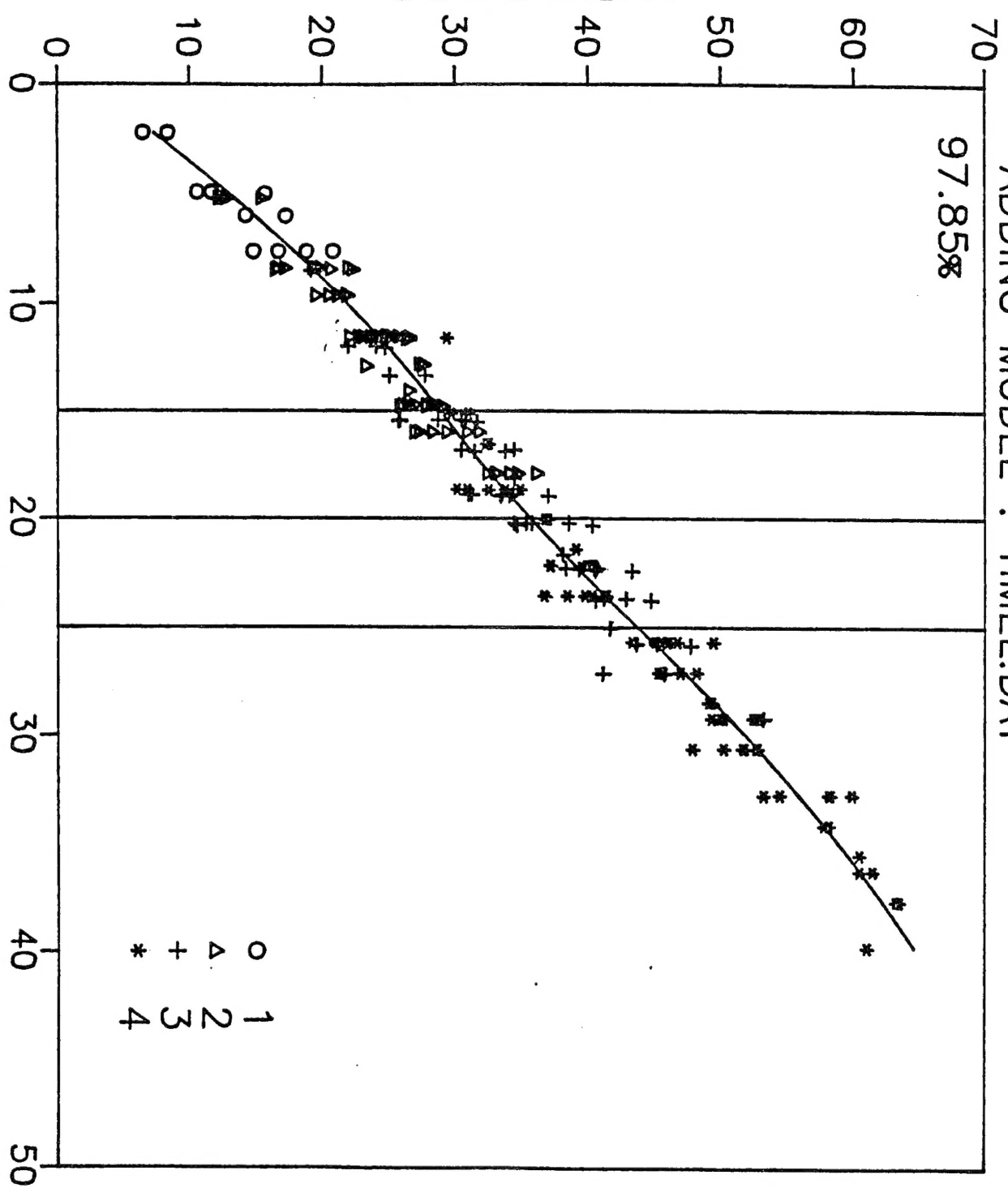


Figure 19a

ADDING MODEL : TIMEL.DAT

97.85%

JUDGMENT



ESTIMATED X
Figure 18b